Active inference: building a new bridge between control theory and embodied cognitive science

Manuel BALTIERI

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the

School of Engineering and Informatics
University of Sussex
January, 2019

Declaration of Authorship

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signed:

Date:

"[T]he rule "collect truth for truth's sake" may be justified when the truth is unchanging; but when the system is not completely isolated from its surroundings, and is undergoing secular changes, the collection of truth is futile, for it will not keep. "

Ashby (1958)

Abstract

The application of Bayesian techniques to the study and computational modelling of biological systems is one of the most remarkable advances in the natural and cognitive sciences over the last 50 years. More recently, it has been proposed that Bayesian frameworks are not only useful for building descriptive models of biological functions, but that living systems themselves can be seen as Bayesian (inference) machines. On this view, the statistical tools more traditionally used to account for data in biology, neuroscience and psychology, are now used to model the mechanisms underlying functions and properties of living systems as if the systems themselves were the ones "calculating" those probabilities following Bayesian inference schemes. The free energy principle (FEP) is a framework proposed in light of this paradigm shift, advocating the minimisation of variational free energy, a proxy for sensory surprisal, as a general computational principle for biological systems. More intuitively and under some simplifying assumptions, the minimisation of variational free energy reduces, for an agent, to the minimisation of prediction errors on sensory input. Initially proposed as a candidate unifying theory of brain functioning, the FEP was later extended to encompass hypotheses on the origins of life, and is nowadays discussed in the cognitive science community for its possible implications for theories of the mind. In particular, one of the most popular process theories derived from the FEP, active inference, describes a biologically plausible algorithmic implementation of this principle with several repercussions on our understanding of cognition.

In this thesis, I will focus on the role of this process theory for action and perception. In active inference, the two of them are combined in a closed sensorimotor loop as co-dependent processes of minimisation of a single loss function, variational free energy, with respect to different sets of variables. Building on this, I will suggest that some of the core ideas of active inference are best seen in terms of enactive, embodied, extended and embedded (4E) theories, in contrast to the majority of the literature emphasising its apparent connections to more traditional, computational, accounts of the mind. In particular, I will develop this argument by focusing on some proposals central to 4E approaches: (a) the non-brain-centric nature of cognitive processes, (b) the lack of explicit representations of the world, (c) the coupling of agent-environment systems and (d) the necessity of real-time feedback signals from the environment. Under the FEP formulation, I will present a series of case studies with mainly two objectives in mind: 1) to conceptually analyse and reframe these 4E ideas in the context of active inference, arguing for the advantages of their formalisation in a more general probabilistic (Bayesian) framework and, 2) to present new mathematical models and agent-based implementations of some of the conceptual connections between Bayesian inference frameworks and 4E proposals, largely missing in the literature.

Acknowledgements

I am grateful to my supervisor, Christopher Buckley, for his guidance over the last three years, inspiring me when it wasn't clear where this project was going and providing me with accurate feedback when things finally began to work. He showed me how to be more critical and more constructive. Much of what I am as a scientist now, I owe to him. I want to thank Thomas Nowotny, my second supervisor, who lent an ear on several occasions. He regularly helped me with the most disparate issues and made sure I always had a backup plan when the exact direction of this thesis was still unclear. A special mention to my examiners, Andy Clark and Daniel Polani, whose input surely improved the presentation of this work.

I thank my colleagues at Sussex: Simon McGregor for our discussions on information theory, probability theory and cognitive science and Chris Thornton for his careful comments on some of my caricatured descriptions of different ideas in cognitive science. I am also greatly indebted to colleagues from the Sackler Centre for Consciousness Science: Keisuke Suzuki for our conversations on embodied cognitive science, dynamical systems and physics, Warrick Roseboom for pointing out some of my naive beliefs and assumptions regarding empirical (neuro)science, Lionel Barnett for the invaluable time he spent covering different aspects of stochastic processes and time series analysis, Anil Seth for sparking my initial interest in the theories presented in this thesis, for encouraging me to interact with his group and for providing support in different ways during this journey.

The time spent in Tokyo at EON/ELSI between 2017-18 during an intermission period was invaluable and for this opportunity I must thank Olaf Witkowski for believing in me in the first place. Thanks to him and to Nathaniel Virgo, I had the chance to work on some incredibly interesting research questions while surrounded by wonderful colleagues. This experience was one of the most rewarding in my (short) career, widening my perspectives on fundamental questions of science. I wish to thank in particular Martin Biehl, Nicholas Guttenberg, Takuya Isomura and Lana Sinapayen for the conversations we had in Tokyo and around the world. For these conversations I also want to thank Taro Toyoizumi, Hideaki Shimazaki, Takashi Ikegami and Ryota Kanai who hosted me in their groups on different occasions to discuss my ideas, anything and everything. Taro and Hideaki especially helped me while I was trying to sort out some of the mathematical details presented in later chapters, thanks to them I came back to Brighton knowing that my ideas could work.

My office-mates also played an important role at different stages of this project. Esin Yavuz helped me when I first settled in the office and listened to the crazy ideas of a first year PhD student, giving me good advice on several important matters. Mario Pannunzi forced me to consider ideas beyond maths and suggested ways to clarify some of the points presented in early chapters.

To my family and to the friends who supported me during this long journey, you have my deepest gratitude.

Contents

List of Figures

List of Tables

To my late father

Chapter 1

Introduction

A deep paradigm shift is affecting studies of biological and cognitive systems: the popular Bayesian frameworks previously used to only descriptively model natural phenomena, are now proposed as mechanistic explanations of the systems themselves. This conceptual move is quite radical: from applying inferential models to datasets without necessarily investigating the underlying mechanisms generating such observations, to thinking about agents as the ones performing inference on their sensory stimuli in order to handle uncertain environments and noisy information. In this light, Bayesian inference processes are thought to be implemented by the agents themselves with scientists simply describing how agents naturally account for their incoming sensations.

The Bayesian brain hypothesis, predictive coding, the free energy principle and active inference are increasingly popular ideas describing biological functions that claim to unify our understanding of life and cognition within a general mathematical framework derived from information and control theory, statistical physics and machine learning (Dayan et al., 1995; Rao and Ballard, 1999; Knill and Pouget, 2004; Friston, Kilner, and Harrison, 2006; Clark, 2013; Hohwy, 2013; Bogacz, 2017; Buckley et al., 2017). The free energy principle (FEP), in particular, is proposed as a unifying framework for the natural and cognitive sciences, with many of the other theories and ideas emerging as special cases/corollaries (Friston, 2010b). According to the FEP, living systems must minimise the surprisal (Tribus, 1961), or information content/self-information, of their sensory states (Friston, Kilner, and Harrison, 2006; Friston, 2010b; Friston, 2012), while meeting their normative constraints (Friston, Thornton, and Clark, 2012; Barto, Mirolli, and Baldassarre, 2013; Clark, 2013). This surprisal is an information theoretical quantity measuring how implausible a state is: frequent states have low surprisal while less frequent ones have high surprisal. Normativity is invoked to describe *value* together with this notion of frequency, which ought to be based on the specific needs of an agent (Friston, Thornton, and Clark, 2012; Barto, Mirolli, and Baldassarre, 2013). For instance, low surprisal states for fish, e.g., to be in water, should not have the same information content for birds. Following this, biological systems are described tautologically as the ones that can frequently occupy a set of states that allow them to exist. Surprisal minimisation, as we will see more in detail in the following chapters, cannot be directly implemented in practice due to the intractability of integrals over continuous spaces of sensory inputs and their hidden states in the world. A possible approximation entails the use of variational free energy to provide a proxy for surprisal: by minimising a known quantity, variational free energy (hence the name: *free energy* principle), a system can effectively reduce the surprisal of its sensations. Active inference (Friston, Daunizeau, and Kiebel, 2009; Friston et al., 2010a), an extension of predictive coding models of the cortex (Rao and Ballard, 1999; Lee and Mumford, 2003; Spratling, 2017), rests on the minimisation of variational free energy, implementing an optimisation scheme that defines a set of processes suggested as possible descriptions of perception, action, learning and attention among others (Friston, Kilner, and Harrison, 2006; Friston et al., 2010a; Feldman and Friston, 2010; Friston, 2010b).

Most of the proposals advanced by the FEP have insofar focused on brain sciences with a structured framework based on Marr's architecture (Marr, 1982) (see Fig. 1.1), outlining why, how and what brains exactly ought to do (Friston, 2008a; Parr and Friston, 2018c). The FEP, however, is proposed to describe and provide new interpretations on several aspects of living creatures beyond brains (Friston, 2012; Friston, 2013), and therefore Marr's architecture is, in this case, adopted as a frame of reference for models of biological systems more in general. In this framework, the minimisation of surprisal (through the minimisation of variational free energy) stands at the computational level (Friston, 2010b; Friston, 2012; Friston, 2013). Process theories such as predictive coding (Rao and Ballard, 1999; Friston and Kiebel, 2009c) and active inference (Friston, Daunizeau, and Kiebel, 2009; Friston et al., 2010a), based on models such as the Laplace (neural) code (Friston, 2008a; Friston et al., 2012; Buckley et al., 2017) are thought to be, respectively, algorithms and encodings that can perform surprisal minimisation in a biological plausible way (Aitchison and Lengyel, 2017). Some proposals on the specific neural mechanisms implementing these algorithms at the lowest level, specifically in terms of neural architectures, have then been advanced for instance in Friston (2008a), Friston and Kiebel (2009b), Bastos et al. (2012), Isomura, Kotani, and Jimbo (2015), Isomura and Friston (2018), and Keller and Mrsic-Flogel (2018), but the exact mechanisms remain unclear. This thesis focuses on a combination of the computational and algorithmic levels just described, and in particular on new interpretations of perception and action suggested by active inference.

The origins of the free energy principle can be traced back to a set of ideas developed in a variety of different fields. One of the most popular perspectives associates the FEP to the notion of "unconscious inference" developed by Helmholtz in the 19th century (Helmholtz, 1867), however see Bruineberg, Kiverstein, and Rietveld (2018). On this view, perception is best thought of as the inference of the most likely causes for a particular observed stimulus via a combination of top-down generated hypotheses and bottom-up error correction processes (Gregory, 1980; Lee and Mumford, 2003). This is then described in more detail using models of predictive coding in the cortex (Rao and Ballard, 1999; Friston, 2005a; Spratling, 2017), equivalent to

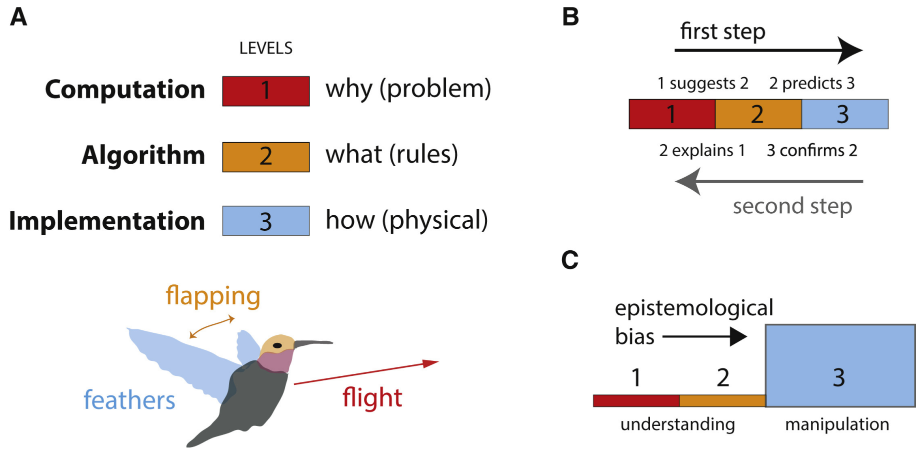


FIGURE 1.1: **Marr's levels of analysis.** In his celebrated architecture, Marr proposed to study a system implementing different levels of analysis. In A) we see an example, studying an organism requires an understanding of the reasons behind its behaviour (Q: Why does it behave this way? A: To fly), the rules specifying the way it behaves (Q: What does it do to fly? A: It flaps its wings) and the mechanisms by which these rules are implemented in the physical world (Q: How does the flapping work? A: Via the use of feathers). In B) we see that these processes are however not disconnected and a complete model should include several cycles of investigation including more specific constraints emerging from the connection of different questions. In C), we report the trend in the neurosciences, as highlighted by the authors of Krakauer et al. (2017).
Reproduced from (Krakauer et al., 2017) with permission from Elsevier Inc.

surprisal minimisation under a set of biologically plausible assumptions (Friston and Kiebel, 2009c). Following this idea, accounts of action (Friston, Daunizeau, and Kiebel, 2009; Friston et al., 2010a; Friston, 2011) and other cognitive functions (Friston, 2010b; Feldman and Friston, 2010) are derived as extensions of the core scheme based on Helmholtzian ideas of (unconscious) inference.

A second proposal outlined by Seth (2014b), suggests that the roots of the FEP can (also) be found in last-century cybernetics approaches to the study of living systems, and in particular brains (Ashby, 1957; Wiener, 1961). This view, providing a fundamental shift in the interpretation of the FEP, is especially relevant for the themes developed in this thesis. Following Seth, in fact, the FEP does not describe how an agent builds accurate and precise world models, but rather a methodology focusing on ideas of homeostasis and control (Ashby, 1960), teleology (Rosenblueth, Wiener, and Bigelow, 1943), autopoiesis (Maturana and Varela, 1980) and ecological psychology (Gibson, 1979). In particular, homeostasis is here used in the Ashbyan sense of maintenance of "essential variables" within boundaries for the survival of an agent (Ashby, 1960). Teleological models (i.e., models defining the purpose of a system) are loosely based on the idea of regulation against external disturbances from the environment via negative feedback mechanisms (Rosenblueth, Wiener, and Bigelow,

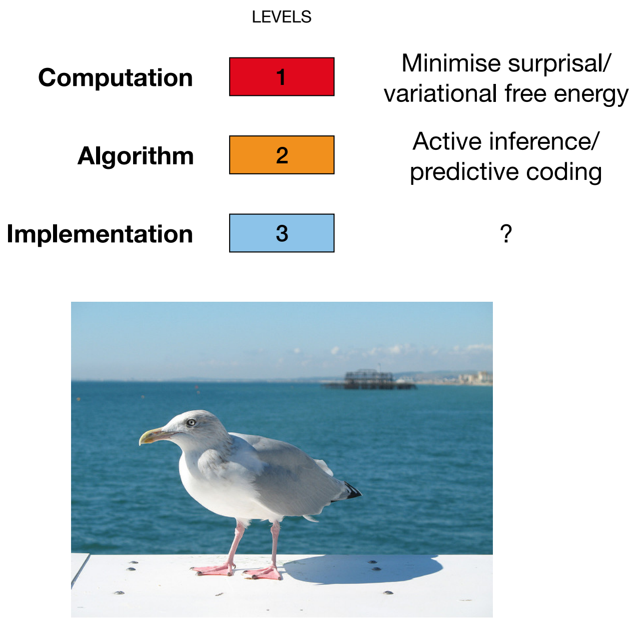


FIGURE 1.2: **Marr's levels of analysis for the FEP.** The computational levels entails the minimisation of sensory surprisal, approximated by the minimisation of variational free energy. The most prominent proposals of an algorithmic or process theory are active inference and predictive coding. At the implementation level we find mostly "evidence consistent with" predictive coding architectures but clear evidence is still lacking (see also Chapter 8).
The image is reproduced from nataliej. "Seagull 5". 23 September 2006. Online image. Flikr. 8 January 2019. www.flickr.com/photos/nataliejohnson/252782639/.

1943) that require an agent to be a good model of such disturbances (Ashby, 1957; Conant and Ashby, 1970; Francis and Wonham, 1976; Sontag, 2003). These two ideas are deeply intertwined and revolve around a different notion of "world model", a model inspired by Conant and Ashby, 1970 and only meant to guarantee survival through homeostatic regulation, not necessarily carrying any objective description of the physical world useful to anyone but the agent itself. As suggested in Bruineberg, Kiverstein, and Rietveld (2018), the FEP is also clearly inspired by ideas of autopoiesis (Maturana and Varela, 1980) as shown in Friston (2013), and affordances in ecological psychology (Gibson, 1979), see for instance Linson et al. (2018). The former highlights possible connections to 4E theories of cognition, down to the origins of life (Friston, 2013; Friston et al., 2015). The latter sees Gibsonian affordances as a natural source of inspiration for ideas emphasising agent-environment coupling, and concepts such as Umwelt (Clark, 1998) and niche construction (Odling-Smee, Laland, and Feldman, 2003; Bruineberg et al., 2018).

Historically, the cognitive science community has adopted different mathematical formalisms for the study of cognition. Often these formalisms have also been proposed as metaphors of the underlying cognitive processes themselves: information theory for computational/representational accounts (McCulloch and Pitts, 1943; Marr, 1982; Fodor, 1983) and dynamical systems theory for work in 4E (Enactive, Embodied, Extended and Embedded) theories (Braitenberg, 1986; Brooks, 1991a; Beer, 1997; Clark, 1998; Van Gelder, 1998). Traditionally, computational theories have paid little attention to closed loop dynamics in agent-environment coupled systems (Clark, 1998), while dynamical systems proponents have not generally tackled more "cognitive-demanding" problems generalising reflex-based behaviour using mappings of statistical regularities (Clark and Thornton, 1997). In this thesis, I will argue that the FEP framework provides a unifying perspective of dynamical systems and information/probability theory for the study of sensorimotor loops in agents, complementing and possibly extending work such as Ashby (1958), Tishby and Polani (2011), and Beer and Williams (2015). One of core themes driving this proposal is what I consider a profound misconception about the use of Bayesian inference frameworks for the study of cognitive systems: the idea of inference to the best explanation and its connections to Bayesian confirmation theory (Douven, 2017) and the notion of analysis by synthesis (Neisser, 1967; Gregory, 1970; Dayan et al., 1995; Yuille and Kersten, 2006). According to this view, inferential methods are primarily processes of knowledge discovery. In Bayesian accounts of cognition, agents are thus systems whose main purpose is to collect information from the world in order to explain their sensations. This perspective is usually adopted in models of what I will call *passive agents* from now on, systems whose goal is to generate an objective understanding of their environment. The FEP and one of its related proposals in particular, predictive coding (Rao, 1999; Friston and Kiebel, 2009c), are often introduced using the idea of passive agents fitting the more traditional, computational interpretation of cognitive systems as having complex and precise world models (Rao and Ballard, 1999; Friston, Kilner, and Harrison, 2006; Hohwy, 2013). I will argue that active inference on the other hand is more in line with 4E theories and with a definition of agents as *acting* on the world rather *explaining* the world (Barandiaran, Di Paolo, and Rohde, 2009; McGregor, 2017; Biehl, 2017). I will explain how the two ideas of passive and active agents can, to some level, co-exist and sometimes be confused due to the attempts made by the FEP to unify different (but dual, as explained in Kalman (1960b), Mitter and Newton (2003), Kappen (2005), Todorov (2006), Todorov (2008), and Todorov (2009b)) mathematical formalisms used for the study of cognitive processes. On the one hand, the more dynamical-systems-oriented view of (optimal) control theory accounts of action and on the other, the Bayesian inference schemes adopted as models of perception. In the FEP formulation, it will become clear that the more intuitive idea of inference to the best explanation is just one side of the coin, and in particular that the teleological and normative interpretations of agents' behaviour inspired by ideas such as autopoiesis and homeostasis are consistent with active inference, providing thus a more natural venue for studies of agency and behaviour.

1.1 Thesis contributions

With a few exceptions (Friston, Daunizeau, and Kiebel, 2009; Schwartenbeck et al., 2015), existing implementations of the FEP have insofar focused on a very specific notion of generative models for passive agents: models that represent and mirror the environment, a representational stand-in for the physical world within the agent. It has been suggested that this need not to be the case (Clark, 2015b; Clark, 2015a), and that the FEP could be interpreted following 4E accounts of cognition (Clark, 2015b; Pezzulo et al., 2017; Bruineberg, Kiverstein, and Rietveld, 2018; Kirchhoff, 2018; Allen and Friston, 2018). These suggestions, however, have provided no direct computational implementations supporting the connections between the FEP and core ideas of 4E cognition. In this light, the aim of this thesis is two-fold. On a more conceptual level, I will discuss some specific connections between approximate Bayesian inference/optimal control frameworks and 4E perspectives for the study of cognition. My arguments will be based mainly on hypotheses regarding perception and action as described by (Bayesian) inference and (optimal) control respectively. The majority of 4E proposals are still based on frameworks relying on (often small) deterministic dynamical systems that hardly scale up to capture more complex features of biological or even artificial agents (Kirsh, 1991; Brooks, 1997). On the other hand, the use of probabilistic and information theoretical frameworks shows promising results for the study of different systems including noise (Longtin, 2003; Todorov, 2004; Faisal, Selen, and Wolpert, 2008), uncertainty (Griffiths et al., 2010; Engel, Friston, and Kragic, 2016) and ambiguity (Knill and Pouget, 2004), but often lacks any grounding of ideas such as embodiment, situatedness and coupling with the environment (Newen, De Bruin, and Gallagher, 2018).

The role of probabilistic and information theoretical frameworks for the study of the mind constitutes a controversial topic in cognitive science, with work suggesting that approaches such as the FEP are aligned with traditional, representationalist views of cognition (Froese and Ikegami, 2013; Gładziejewski, 2016). Part of the cognitive science community does not see this as a problem (Hohwy, 2013), while others claim it is its main drawback (Gładziejewski, 2016; Zahavi, 2017). Others have argued that the FEP is more consistent with a 4E perspectives of cognition, claiming that its strengths reside in generative models with no explicit representational features (Bruineberg, Kiverstein, and Rietveld, 2018). A different perspective highlights the potential of theories such as the FEP for the formalisation of "action-oriented" views of cognition (Clark, 2015a; Clark, 2015b; Engel, Friston, and Kragic, 2016; Pezzulo et al., 2017; Allen and Friston, 2018), attempting to reconcile traditional views and 4E theories on the ground of action as a central process for adaptive systems. My work will support 4E views, suggesting at the same time that information and probability theory can be useful frameworks to treat noise and uncertainty, without necessarily bringing in representationalist arguments typical of cognitivist descriptions of the mind.

Making use of the mathematical relation between information theoretical and dynamical/control-theoretical perspectives on cognitive processes, I will then provide models supporting the connections between the FEP and 4E theories by focusing on definitions of sensorimotor loops where uncertainty and noise are key features of the agent-environment dynamics. The implications of this relation remain vastly unexplored with a few exceptions found for instance in Tishby and Polani (2011) and Beer and Williams (2015). I will thus propose examples where the integration of these two mathematical formalisms sheds light on new interpretations of cognitive systems, for instance the idea of perception and action as deeply entangled functions of embodied and situated agents (Clark, 1998; Wilson, 2002; Beer and Williams, 2015; Di Paolo, Buhrmann, and Barandiaran, 2017). My contribution is set out to extend the FEP to studies of sensorimotor loops in a 4E spirit, with new concrete proposals on issues arising in the 4E literature regarding the use of information theoretical tools. In particular, in Chapter 4 I will provide an initial description of the simplest sensorimotor loop under active inference to introduce the core ideas developed in later chapters. In Chapter 5 I will show how the FEP can formalise ideas such as the importance of the dynamical coupling between an agent and its environment. Homeostatic regulation with minimal models of world dynamics will then be introduced in Chapter 6. Finally, the non-modular nature of action and perception will be discussed in Chapter 7.

1.1.1 Published work

Most of the work presented in this thesis was either published, submitted for publication or is currently in preparation for publication in collaboration with Christopher Buckley. The list of manuscripts includes:

- "An active inference implementation of phototaxis" In: Proc. Eur. Conf. on Artificial Life. 2017 (Baltieri and Buckley, 2017), presented in Chapter 5 with a corrected notation
- "A probabilistic interpretation of PID controllers using active inference" In: From Animals to Animats 15., 2018 (Baltieri and Buckley, 2018a), presented and extended (see below) in Chapter 6
- "The modularity of action and perception revisited using control theory and active inference" In: The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE) (Baltieri and Buckley, 2018b), presented and extended (see below) in Chapter 7.
- "Nonmodular architectures of cognitive systems based on active inference" (Baltieri and Buckley, 2019b), submitted to IJCNN (International Joint Conference on Neural Networks), presented and extended (see below) in Chapter 7

- "PID control as a process of active inference with linear generative models" (Baltieri and Buckley, 2019c), expanding on Baltieri and Buckley (2018a) and presented in Chapter 6, submitted for journal publication and now under review
- "Modularity, the separation principle and active inference" (Baltieri and Buckley, 2019a), expanding on Baltieri and Buckley (2018b) and Baltieri and Buckley (2019b), presented in Chapter 2 and Chapter 7

1.1.2 Limitations

Some of the major limitations will be initially pointed out here and re-elaborated in the concluding chapter. This thesis focuses on formulations of the FEP and active in continuous space and time that were developed mainly until 2011 and followed by implementations of (PO)MDPs (Partially Observable Markov Decision Processes) in discrete time, see for instance (Friston, Samothrakis, and Montague, 2012a; Friston et al., 2013; Friston et al., 2015; Friston et al., 2017). The continuous time time formulation adopted in this work makes use of a set of assumptions that greatly simplify the variational treatment of Bayesian inference problems, see Friston et al. (2007), Friston, Trujillo-Barreto, and Daunizeau (2008), Friston et al. (2010b), and Buckley et al. (2017), while for a more general treatment of variational methods in a Bayesian context see Beal (2003) and Bishop (2006). The more technical aspects will be developed and discussed in the next chapters, however, I will point out some general limitations here.

Firstly, this treatment only covers unimodal probability distributions due to the Laplace assumption (Friston et al., 2007), approximating an unknown distribution using a Gaussian form. This is sometimes introduced as a strength, proposing the Laplace *encoding* as a plausible neural code based on the sufficient statistics of a Gaussian density (Friston, 2009), but this hypothesis is not confirmed empirically. We could extend this framework by using methods that can approximate multi-modal distributions (similar to particle filters) under the same variational formulation (Friston, 2008b), but their simulation is computationally expensive.

Secondly, while many authors (Lee and Mumford, 2003; Friston, 2008a; Hohwy, 2013; Clark, 2013) stress the importance of hierarchical Bayesian models, especially for (deep) models of perceptual processing, my focus is on minimal (1-layer) architectures. It has also been suggested that hierarchical models can combine the continuous and discrete time formulations of the FEP, , (Parr and Friston, 2018c), providing more complete schemes for the treatment of mixed (continuous and discrete) problems thought to arise in biological systems, typically in terms of continuous inputs/outputs and discrete decision making processes. These aspects remain largely unexplored in my work, focusing mostly on minimal (rather than complex and hierarchical) accounts of behaviour.

Lastly, the formulation I use in this thesis is essentially "model free", in a reinforcement learning sense, since agents are already endowed with appropriate heuristics that map actions to suitable outcomes towards a goal (see Chapter 3 and onwards). The agents implemented here can thus minimise a variational free energy functional with no dependences on future outcomes (Friston, Daunizeau, and Kiebel, 2009; Friston, 2010b; Friston, Samothrakis, and Montague, 2012a), cf. the discussion on inference approaches to control and behaviour in Kappen, Gómez, and Opper (2012) and the subsequent POMPD's *expected* free energy treatment proposed in Friston, Samothrakis, and Montague (2012a) and Friston et al. (2015). In this thesis, policies are built as sequences of independent actions obtained through instantaneous minimisations of variational free energy. Considering the complexity of formulating accounts of behaviour with time-dependent policies conditioned on future (therefore hidden) states in continuous time, frameworks proposed since 2012 (Friston, Samothrakis, and Montague, 2012a) in terms of expected (Friston et al., 2015) and generalised free energy (Parr and Friston, 2018a) are not included in my presentation of the FEP. Such proposals include more general accounts of, for instance, planning (Friston, Samothrakis, and Montague, 2012a; Friston et al., 2015)) and behaviour based on (epistemic) uncertainty minimisation (Friston et al., 2016a; Parr and Friston, 2017) that are not (yet) developed for the continuous time formulation of the FEP and thus will not be covered.

Chapter 2

Background

In this chapter I will briefly overview the 4E ideas central to this thesis. I will then focus on the relationship between 4E theories and the cybernetics movement, in particular on some of the approaches proposed for the study of homeostatic regulation, attempting to unify information and control theoretical notions of entropy and control under a more general mathematical theory with implications for the biological and cognitive sciences. This will (re)connect some of the mathematical methodologies adopted by 4E approaches for the study of agents, and based on dynamical systems theory, to ideas of goal-directedness and teleology in a control- and information-theoretic sense. To explore the FEP's proposal to unify information, control and studies of natural systems under the 4E frameworks, I will specifically cover action-perception loops. After discussing the most recent theories of perception and action in the natural and cognitive sciences, including for instance the Bayesian brain hypothesis, predictive coding/processing and optimal control, I will highlight their individual differences and their conceptual connections to theories of *estimation* and *control*. Finally, I will introduce the free energy principle, bringing together these theories in a single framework. I will then discuss its most popular implementations, highlighting an extensive lack in the literature of models showing the possible connections between Bayesian/control theory approaches and 4E cognition. A mathematical formulation of the FEP will then be the focus of the next chapter.

2.1 Embodied, enactive, extended and embedded (4E) cognition

In the last few decades, the cognitive sciences have seen an increasing number of approaches questioning the "classical sandwich" (Hurley, 2001) of the mind, the traditional proposal that perception, higher order cognitive functions such as modelling, planning and memory, and action are implemented as a series of activities in a fundamentally sequential, feedforward fashion (Fodor, 1983), (see also Cisek (1999) for a short but complete introduction to these topics). Many of these approaches implement ideas inspired by literature on 4E cognition, see for instance Varela, Thompson, and Rosch (1991), Clark (1998), Pfeifer and Scheier (2001), Gallagher (2006), Di Paolo, Buhrmann, and Barandiaran (2017), and Newen, De Bruin, and Gallagher (2018) for extensive treatments and reviews. In an attempt to capture some of the central arguments against the cognitivist approaches to the study of the mind, different areas are included in the "4E" definition: embodied, enactive, extended and embedded theories. It is important to highlight that differences, and perhaps some incompatibilities among these areas do exist, for instance on how strongly traditional ideas of representations and computation are rejected or effectively (mildly) accepted (Newen, De Bruin, and Gallagher, 2018). Crucially, however, all 4E approaches share an attempt to abandon the "Cartesian theatre" (Dennett, 1993) metaphor of the mind: a computational, brain centric allegory of mental processing that, 4E practitioners suggest, needs to be re-evaluated under a different light (Newen, De Bruin, and Gallagher, 2018). Some of the common principles in 4E theories include embodiment, situatedness, feedback from the environment and a non-modular perspective of perception, cognition and action (Clark, 1998; Pfeifer and Scheier, 2001; Wilson, 2002). These ideas have been extensively explored in different research areas, including for instance robotics, artificial intelligence, philosophy of mind, psychology, anthropology, linguistics and artificial life, see Boden (2006) for an historical account of their impact on different areas of cognitive science.

In this thesis, I will not attempt to give a full overview of 4E theories, but rather focus on the mathematical treatments and implementations of some of their central tenets, starting from how sensorimotor loops have been modelled in the 4E literature so far, and proposing some new perspectives. The mathematical underpinnings of 4E approaches are heavily reliant on dynamical systems theory, due to a strong emphasis on the dynamical interaction of brain-body-environment architectures (Van Gelder, 1998; Beer, 2008). Some of the most prominent models of agent-environment systems include direct applications of dynamical systems theory (Beer, 1995; Beer, 1997; Beer, 2008), evolutionary robotics (Cliff, Husbands, and Harvey, 1993; Nolfi, Floreano, and Floreano, 2000) and behaviour-based robotics (Brooks, 1986; Brooks, 1991a; Brooks, 1995). Alternative frameworks, and in particular I refer to (Shannon) information theory (Shannon, 1948), are on the other hand often associated to information processing and traditional cognitivist arguments promoting the metaphor of the mind as a machine (Turing, 1937; McCulloch and Pitts, 1943; Boden, 2006). These proposals, dating back to the work of Tuning (Turing, 1950; Boden, 2006), are usually based on an interpretation of the mind as processing information on a Von Neumann architecture (Neumann, 2012), including ideas such as memories, symbols, buffers, logical gates, etc. (McCulloch and Pitts, 1943; Newell, Simon, et al., 1972; Fodor, 1983), and still constitute the dominant framework in the cognitive sciences. 4E research rejects theories based on information processing metaphors, following the idea that classical information theoretical measures, for instance mutual information and entropy as originally defined by Shannon for communication problems (Shannon, 1948), inherently depend on the presence of abstract *symbols*. Furthermore, they are based on atemporal definitions, failing to capture the intrinsically dynamic nature of cognition. Information measures are still, nonetheless, normally used as tools for post-hoc analysis of different models (Pfeifer and Scheier, 2001).

2.1.1 The cybernetics roots of 4E

Many of the ideas in 4E theories are inspired by previous work in cybernetics (Wiener, 1961; Ashby, 1957) and influenced by approaches involving feedback, (ultra)stability, closed-loop dynamics and circular causality (Froese and Stewart, 2010; Froese, 2010; Froese, 2011; Villalobos, 2013), with influential examples such as Walter's tortoises (Walter, 1950) and Braitenberg's vehicles (Braitenberg, 1986) (see however Maturana (2011) for ideas claimed to be explicitly *not* inspired by cybernetics). In the cybernetics movement, feedback loops and homeostatic control in agent-environment systems constituted a central area of research. Based on ideas of homeostatic control, Ashby developed his "law of requisite variety" (Ashby, 1958), claiming that to stabilise a system, a controller must have access to a number of states (at least) equal to the number of states of the system to be controlled. In his work, then, Ashby explicitly attempted to relate concepts of control and Shannon information theory under the same mathematical formulation (Ashby, 1958). These efforts culminated in the good regulator theorem (Conant and Ashby, 1970), claiming that *"every good regulator of a system must be a model of that system"*: a controller must be a model (in some broad sense) of the system being regulated. Since then, however, many of these ideas have taken off independently (Boden, 2006). Information theory became a central focus in the field that would then be called artificial intelligence, especially GOFAI (Russell and Norvig, 2009). Control theory contributed, often indirectly, to the development of ideas related to intrinsic homeostatic regulation, e.g., autopoiesis (Maturana and Varela, 1980), relying on concepts of dynamical systems without extrinsic value functions (however see once again Maturana (2011) and Maturana's stance on formal mathematical definitions). In parallel, control theory has also developed into modern *optimal* control theory (Anderson and Moore, 1990; Stengel, 1994) based on Pontryagin (Pontryagin et al., 1962) and Bellman (Bellman, 1957b) formulations.

Control theory and information theory, heavily used by cybernetics researchers in the last century, are nowadays however less popular in 4E models, often associating the definition of computational goals (cf. Marr's architecture (Marr, 1982) and Chapter 1) of a system to more traditional, symbolic accounts of cognitive science, see Di Paolo (2005) and Villalobos (2013) for thorough discussions. In particular, some of the conceptual inconsistencies between dynamical systems approaches to cognition and optimal control accounts of behaviour often emerge from the use of *extrinsically* defined value functions advocated in the latter (see Todorov (2006) and Schaal, Mohajerian, and Ijspeert (2007) for discussions on this and related topics). The presence of goals defined externally by an experimenter works against intrinsic ideas of autonomy, adaptivity, individuation and in particular autopoiesis as originally formulated in Maturana and Varela (1980), rejecting the teleological approach described by Rosenblueth, Wiener, and Bigelow (1943). Following the original definition of autopoietic accounts of life and cognition, teleological frameworks (including information and control theory) are seen as fundamentally symbolic, with a functional role that can only be attributed by an external observer (Di Paolo, 2005). However the debate remains unsettled (Villalobos, 2013), especially considering more recent descriptions of autopoiesis as "embodied teleology" (Weber and Varela, 2002) by one of the original contributors, Varela, (together with Weber), a view not shared by the other original author, Maturana, see for instance Maturana (2011)). This definition can reintroduce the use of information and control metaphors for the study of living systems in a 4E context, as argued in Villalobos (2013). In strong support of the reconnection between cybernetics ideas and 4E cognition, the FEP proposes also to replace extrinsic cost and value functions with intrinsic priors. On this view, priors encode normative drives imposed by the minimisation of sensory surprisal in order for an agent to exist, consistent with accounts of autopoiesis (Friston, 2013).

2.1.2 4E cognition and models of the environment

More recently, some of the ideas found in cybernetics, attempting to unify control and information theory, have re-emerged in the field of stochastic optimal control, where the problems of estimation and control of a system have been shown to share a common mathematical background (Anderson and Moore, 1990; Stengel, 1994; Todorov, 2008; Kappen, Gómez, and Opper, 2012). This unification is thought to play an important role in the biological sciences, see for instance Sontag (2003). In the study of living systems, while a sharp separation between a controller and the system to control is not as clear, concepts such as the good regulator theorem and its direct and more formal extension, the internal model principle (Francis and Wonham, 1976), are becoming established modelling frameworks. These methodologies have been applied to the study of systems such as bacteria, e.g., E. Coli, where mechanisms of *perfect adaptation* explaining chemotaxis have been found in the bacteria's signal transduction pathway (Yi et al., 2000). Perfect adaptation, a popular term in cellular and molecular biology, describes the ability of living systems and biochemical networks to adapt their response to the presence of external stimuli by returning to a level of activity near the pre-stimulus baseline. Yi et al. (2000) showed that this mechanism is equivalent to integral control (Åström, 1995), while Andrews, Yi, and Iglesias (2006) used a Kalman filter, a classical estimation method, to determine the properties of this "controller", providing results remarkably consistent with experimental data.

The revival of ideas on the unification of information and control theory has also produced several new proposals for the study of sensorimotor loops more in general, including for instance Ay et al. (2008), Tishby and Polani (2011), and Jung, Polani, and Stone (2011). These results define new information theoretical measures for agents-environment systems over temporal "channels", considering possible dynamical correlations between variables of a agent-environment coupled system over time. All these proposals elegantly deal with one of the main drawbacks pointed out by 4E approaches to cognition for information theory: the atemporality of information measures defined for, initially, spatial channels.

Most of these ideas seem to clash with positions advocated by 4E research, especially the notion of a (possibly) symbolic "internal model" in the internal model principle (Francis and Wonham, 1976) and in the good regulator theorem (Conant and Ashby, 1970). While perhaps not solving the problem entirely, in this thesis I will make an explicit distinction to define operationally the use I make of the words "internal" and "model". This distinction is based on the idea of a system either *having* a model of the environment or a system *being* a model of the environment, see also Seth and Tsakiris (2018) for a similar discussion. *Having* a model is the idea that I associate to cognitivist accounts of the mind, a Cartesian theatre perspective where a "model" exists as an independent and often abstract set of symbols that a system manipulates to act on its environment. *Being* a model, on the other hand, highlights the fact that a system's very existence, properties and functions are in some way isomorphic to the environment (Conant and Ashby, 1970; Sontag, 2003; Friston, 2013) and can be described by a scientist using a set of mathematical objects, i.e., a model. I will avoid, as much as possible, the more confusing use of "internal". In this thesis I will focus on notions of systems *being* a model of their environment, leaving interesting but more problematic developments often associated to ideas of *having* a model, such as counterfactual thinking, allostasis, etc. (Seth and Tsakiris, 2018), for future work. This simplification follows from my focus on minimal cognition and action-perception loops with no higher order functions (although see Chapter 5 for the implementation of a process thought to be close to attention (Feldman and Friston, 2010; Brown et al., 2013)). This choice is also consistent with the initial assumption I made with regards to the study of simple behaviour based on homeostatic control with policies not depending on an agent's future actions and observations (see section 1.1.2).

2.2 Perception as inference (estimation)

An increasingly popular view in the cognitive sciences posits that brains are systems performing (Bayesian) inference and that perception is a process that can be described using methods from classical estimation theory (Knill and Richards, 1996; Knill and Pouget, 2004; Friston, 2005a; Friston, 2010b; Hohwy, 2013; Clark, 2015b).

The main goal of estimation processes is to infer (or estimate) the hidden dynamics of a system given a set of observations that are in general noisy or encoding uncertainty, and a model of the underlying dynamics that also includes noise/uncertainty. The idea of estimation can be historically traced back to Gauss and his method of least squares used to infer parameters of the motion of celestial bodies in studies of astronomy (Sorenson, 1970). Some tremendous advances came in the last century, with the contributions of Fisher, i.e., maximum likelihood estimation, and Kolmogorov and Wiener, see again Sorenson (1970) for a review. A few years later, the emergence of computing machines accelerated the popularisation of Kalman filters (Kalman, 1960b) as an iterative, computationally efficient extension of previous methods (e.g., the Wiener filter, see Chen (2003)). At the same time, this digital revolution promoted the use of new estimation methods based on large-scale sampling techniques, i.e., Monte Carlo methods, to estimate static solutions of different problems. In the 90's, these methods were then extended to sequential estimations of dynamical systems, e.g., particle filters (for a review on sampling methods for estimation see Chen (2003)).

In cognitive (neuro)science, the first connection with estimation theory can be traced back to Helmholtz and his "unconscious inference" hypothesis depicting brains as inference machines (Helmholtz, 1867). This hypothesis later inspired work on the idea of perception as hypothesis testing (Neisser, 1967; Gregory, 1980) or "analysis by synthesis", for a recent review see Yuille and Kersten (2006). On this view, the emphasis of perceptual processes shifts from a purely bottom-up perspective of features extraction (Marr, 1982) to a combination of bottom-up and top-down, generative streams of information combined in a cohesive process of error correction and prediction update. A few decades later, this conceptual treatment was grounded in models of the visual cortex, i.e., predictive coding (Rao and Ballard, 1999; Lee and Mumford, 2003). In this more explicit hypothesis, brains are thought to infer the most likely causes of their sensory data, or rather to estimate the associations between unknown, latent or hidden variables and some observations[1]. These variables are considered to be inherently ambiguous and uncertain: different circumstances can, in fact, generate the same sensory input. In machine learning, the same set of ideas was used in a now classical example, the Helmholtz machine (Dayan et al., 1995). This connectionist method implements two different components, a recognition model and a generative model. The former attempts to estimate the most likely hidden variables of the data provided to the network (i.e., the analogous of sensory input for the brain) while the latter generates sequences of observations given the current best estimates of said variables. In later formulations of predictive coding models and based on this formulation, these two models are proposed to correspond to two different types of cortical processing. More specifically, recognition models can be seen as implementing bottom-up processes thought to represent increasingly more complex statistically regularities and patterns in a hierarchical network. Generative models, on the other hand, generate fictitious observations in a top-down fashion starting from a set of current best estimates of world variables (which are

[1]While a vast majority of the literature on predictive coding/free energy principle assumes that Bayesian frameworks allow to infer the causal structure of some given observations, see for example Clark (2013), Hohwy (2014), and Parr and Friston (2018b), in this thesis I adopt a Pearlian stance, objecting to this notion of causality and to the fact that Bayesian frameworks can ever be used for the study of a meaningful notion of causality (Pearl, 2001).

updated over time by the recognition model). The predictions produced via top-down generative processes are then compared to the real sensory input stream and their mismatch error is used as a mechanism of self-supervision to assess the quality of the bottom-up inferred variables and internally generated sensations. The minimisation of this error thus constitutes a way to improve the estimates of incoming data. In the context of predictive coding models in neuroscience, this minimisation is then associated to processes of perception, or perceptual inference.

Many of these ideas fall, nowadays, under umbrella terms such as predictive coding, predictive processing, Bayesian brain hypothesis and the free energy principle. While all of them share some common background, they generally differ in some important aspects. In the remainder of this section I will attempt to clarify the terminology adopted in different research areas while also reviewing some of the most influential work inspired by these ideas. The free energy principle will be the only one treated separately at the end of the chapter, since it constitutes the conceptual and mathematical basis of the following chapters.

2.2.1 The Bayesian Brain hypothesis

The hypothesis that the brain represents the latent variables of an uncertain world in a probabilistic and Bayes optimal manner to guide actions and behaviour has become dominant in the last 20 years. Bayesian methods have, in fact, proved to be extremely powerful in modelling various aspects of brain functions and motor control (Ernst and Banks, 2002; Rao et al., 2002; Knill and Pouget, 2004; Körding and Wolpert, 2006; Doya, 2007; Griffiths, Kemp, and Tenenbaum, 2008; Griffiths et al., 2010; Penny, 2012). The Bayesian brain hypothesis promotes the idea that brain functions may be described using Bayes optimal computation. This computation is thought to implement processes of probabilistic inference, providing an appropriate set of methodologies to describe biological systems that live in an uncertain world perceived through noisy receptors. Bayesian schemes can be seen as implemented on different levels, from spike trains at a neuronal level to coding in neural circuits and up to behaviour at a systemic level (for a review see Doya (2007)). The focus here is on the latter, given its possible interpretations for global theories of cognition.

Bayes' theorem provides, in this context, an elegant and optimal (Robert, 2007) way of explaining inferential processes. To briefly introduce the general concept, I will readapt an example from Knill and Pouget (2004) combined with an idea from Lee and Mumford (2003). Suppose a person is trying to estimate what object $X = x$ (e.g., a glass) she is looking at, given visual cues $V = v$, where x and v represent specific values of random variables X and V, see Fig. 2.1. In Bayesian terms, this is equivalent to finding the conditional probability $p(x|v)$, that can be expressed using Bayes theorem as:

$$p(x|v) = \frac{p(v|x)p(x)}{p(v)} \tag{2.1}$$

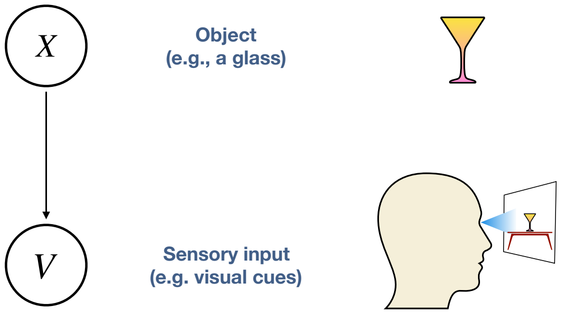


FIGURE 2.1: **A simple example of Bayesian estimation.** A simplified Bayesian network describing sensory input V, e.g., visual cues, (conditionally) dependent on observed objects in the world X, e.g., a glass.

where $p(x|v)$ is called posterior density, $p(v|x)$ is defined as the likelihood of the observation $V = v$ given an object $X = x$ and $p(x)$ is the prior probability of value $X = x$ before observations are taken into account. $p(v)$ is the marginal likelihood or model evidence and can be seen as a normalising factor that ensures that the posterior is expressed on the $[0, 1]$ interval.

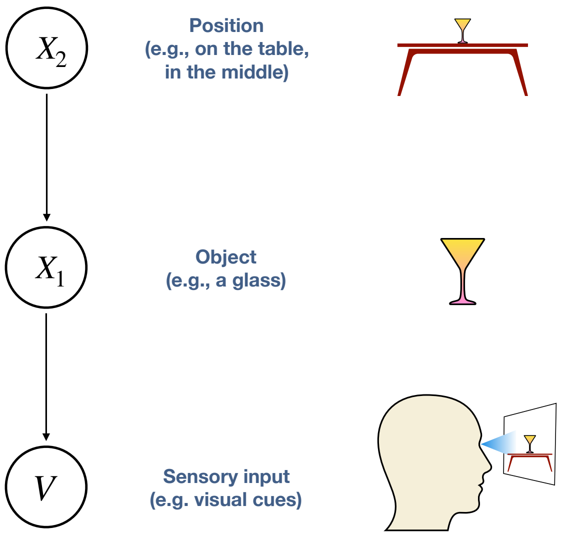


FIGURE 2.2: **A simple example of hierarchical Bayesian estimation.** A simplified Bayesian network with multiple layers describing sensory input V, e.g., visual cues, (conditionally) dependent on observed objects in the world X_1, e.g., a glass, (conditionally) dependent on the position of the objects X_2, e.g., on the table, in the middle.

Several studies reported in Lee and Mumford (2003), Knill and Pouget (2004), Körding (2007), and Griffiths et al. (2010) and references therein suggest that Bayesian optimality criteria may provide relevant insights on the mechanisms driving different accounts of biological functions. The above example, while potentially useful to grasp the general idea behind Bayes theorem, lacks however any hint regarding possible implementations in real systems. In the same way, the "Bayesian brain hypothesis" constitutes a general principle for the investigation of cognitive aspects of living system in light of criteria of Bayesian optimality, but is overall agnostic on the exact algorithmic and implementational details, cf. Marr (1982). In practice, the majority of the models based on this hypothesis relies on hierarchical architectures (Rao and Ballard, 1999; Lee and Mumford, 2003; Friston, 2008a), implementing message passing algorithms over different layers, e.g., belief propagation (Bishop, 2006), expectation propagation (Minka, 2001) or the wake-sleep algorithm (Hinton et al., 1995; Tishby, Pereira, and Bialek, 1999). Hierarchical structures, in this context, lend themselves to implementations where Bayes' theorem is applied between layers under a Markov property, see for instance the previously discussed Helmholtz machine (Dayan et al., 1995). Bottleneck levels (layers with fewer nodes compared to the ones below), in particular, become crucial in this picture since they implement effective inference processes by imposing constraints on data compression (Hinton and Zemel, 1994). As a simple example of hierarchy, although with no real compression, we could extend the previous problem on the estimation of an observed object. In this case, we could include a second, higher layer with a new random variable X_2 that encodes the position of the object represented now by the random variable X_1 (previously X). Subscripts show different layers in the hierarchy, with layer 2 at the top level. In this case, an agent could try to estimate the variable at the highest layer X_2 with a posterior calculated in two steps, see Fig. 2.2. At the first layer of this simple network, the posterior for X_1 becomes

$$p(x_1|v, x_2) = \frac{p(v|x_1, x_2)p(x_1, x_2)}{p(v)} \tag{2.2}$$

Thanks to the Markov assumption that guarantees conditional independence on non-contiguous layers, observations do not depend on all possible higher levels, but just on the one directly above. In other words, we assume that an observation $V = v$ (statistically) depends only on the observed object $X_1 = x_1$, not directly on its position $X_2 = x_2$. Given this assumption, the likelihood reduces to

$$p(v|x_1, x_2) = p(v|x_1) \tag{2.3}$$

The joint prior on X_1, X_2 can be conditioned and rewritten as

$$p(x_1, x_2) = p(x_1|x_2)p(x_2) \tag{2.4}$$

with the prior about the variable on the second layer $p(x_2)$ disappearing if we assume that no a-priori knowledge on the glass' location is available, i.e., we have no a priori knowledge of where it is, i.e., we have a flat prior. The posterior at the first layer then reduces to

$$p(x_1|v, x_2) = \frac{p(v|x_1)p(x_1|x_2)}{p(v)} \tag{2.5}$$

where we initially assumed a flat distribution $p(x_2)$ and where the remaining prior $p(x_1|x_2)$ is equal to likelihood used for the inversion of x_2 at the layer above. Assuming, once again, a Markov property between layers we get in fact

$$p(x_2|x_1) = \frac{p(x_1|x_2)p(x_2)}{p(x_1)} \tag{2.6}$$

With this small example, we can see some of the basic concepts behind hierarchical implementations of the Bayesian computation implied by the Bayesian brain hypothesis. The posterior $p(x_1|v, x_2)$ on layer 1 is constrained by observations $V = v$ from below, and by the best estimates of $X_2 = x_2$ from above. In the simple iterative Bayesian update process proposed here, the posterior $p(x_1|v, x_2)$ becomes the prior $p(x_1|x_2)$ in the next iteration and is used as a constraint for the inference of $X_2 = x_2$ at the layer above, where it represents the likelihood of "observing" variable X_1 before new observations are available. Higher layers thus act as prior constraints on lower ones by providing the best guesses of states at lower levels.

While this scheme offers an elegant implementation of Bayesian schemes, severe limitations arise when considering exact methods for Bayesian inference in more complicated problems. The process is, in fact, often analytically intractable (Bishop, 2006), since the computation of the marginal likelihood (i.e., the normalisation factor in Bayes theorem) for complex and especially continuous random variables is rarely feasible. Many approximations have been proposed in the literature which rely on either deterministic or stochastic methods. Stochastic methods are based on Monte Carlo sampling, see for instance Markov Chain Monte Carlo/particle filtering (Chen, 2003; Bishop, 2006), and while very effective, they need a considerable load of computation that brains and biological systems more in general may not be able to provide, especially in dynamic/fast-paced scenarios (Ashby, 1958)[2]. Deterministic approximations provide a computationally cheap solution, with more biologically plausible implementations. At the same time however, these approximations provide lower quality solutions. Variational methods (MacKay, 2003; Beal, 2003; Bishop, 2006) are among the most popular choices for deterministic approximations and are behind the formulation of the free energy principle (see Chapter 3).

It is often assumed that other terminologies, especially predictive coding and predictive processing, are deeply intertwined if not exact synonyms of the Bayesian brain hypothesis but, as some recent work pointed out, this is more confusing than helpful since Bayesian accounts of cognition are not necessarily tied to the algorithmic implementation proposed by predictive coding, see the separation of computational and algorithmic levels in Chapter 1 and discussions in Thornton (2016) and

[2] See Sanborn and Chater (2016) for a different perspective.

Aitchison and Lengyel (2017).

2.2.2 Predictive coding

Predictive coding (PC) is arguably the most ambiguous term mentioned in the literature, and as shown in Spratling (2017), its exact meaning is clearly context-dependent. Even within the same field (e.g., neuroscience), the terminology is still sometimes confusing due to the different background of some research ideas. Originally developed in signal processing (Makhoul, 1975), PC is a methodology used to define a compression algorithm that allows signal reconstruction with a limited amount of parameters. While the underlying principles still survive in definitions of PC in neuroscience, the signal processing version of predictive coding has roots in linear time series analysis and doesn't provide, for instance, a precision (inverse variance)-weighted mechanism typical of the probabilistic/state-space derivation (Spratling, 2017).

In neuroscience, PC has been used to label different hypotheses. *Dynamic* predictive coding is an application of the original definition derived from signal processing to research on the retina (Srinivasan, Laughlin, and Dubs, 1982; Hosoya, Baccus, and Meister, 2005). Models of dynamic predictive coding mainly focus on physiological implementations of PC in the retina, and their underlying computational principles are subsumed by later models (Huang and Rao, 2011). Rao and Ballard's model (Rao and Ballard, 1999), arguably the most popular theory of PC in neuroscience nowadays, suggests a functional hypothesis of predictive coding mapping to feedforward and feedback connections in the visual system. This model also popularised the idea of "perception as hypothesis testing" (Gregory, 1980) with a biologically plausible algorithmic implementation. Spratling's model (Spratling, 2008) is an adaptation of Rao and Ballard's PC to include ideas of biased competition theories of cortical function, essentially replacing the prediction errors' formulation of Rao and Ballard with a mechanism of modulation (division and multiplication) of new (visual) stimuli. Spratling's work provides a different and interesting angle on mechanisms of (visual) attention, proposing different hypotheses about possible physiological implementation of these models in the brain. Another different implementation related to predictive coding is then proposed in models of Hierarchical Temporal Memory (HTM) (George and Hawkins, 2009; McCall and Franklin, 2013; Kneller and Thornton, 2015; Cowley et al., 2018). HTM constitutes a theory of cortical functioning in the neocortex based on the concept of cortical columns. Insofar, this framework has received less attention in the biological sciences given its focus on artificial neural networks and machine learning applications, emphasising the practical advantages of models claimed to be more biologically plausible. Considering my focus on free energy minimisation schemes and their potential implications for cognitive systems at the computational and algorithmic levels, proposals of different biological implementations will not be considered in the remainder of the thesis. Following the suggestion that the free energy principle (introduced in more detail at the end of this chapter and in the next one) recapitulates previous formulations of predictive coding, and under simplifying assumptions is formally equivalent to the formulation by Rao and Ballard (Friston and Kiebel, 2009c), my work will have however implications for Rao and Ballard's proposal.

According to Rao and Ballard's functional model, the visual cortex combines bottom-up and top-down information processes through feedforward and feedback connections between cortical areas in a hierarchical structure (Felleman and Van, 1991; Mumford, 1992), see also later applications to auditory (Baldeweg, 2006) and the motor cortices (Shipp, Adams, and Friston, 2013; Adams, Shipp, and Friston, 2013). On this view, feedback connections carry top-down predictions from higher cortical areas via a generative model of incoming sensations. Bottom-up connections, on the other hand, convey feedforward prediction errors, i.e., the difference between sensory input and the predictions generated by the cortex, to higher regions in order to update and improve future predictions, cf. the Helmholtz machine (Dayan et al., 1995). Similarly to the more traditional definition of predictive coding in signal processing, in Rao and Ballard's model the (sensory) input is never actually delivered to higher areas. These areas only receive prediction errors, representing the residual/unaccounted/unexplained data, which are used to improve future predictions in order to minimise these errors. Following this model, it has then been suggested that perception (or perceptual inference) is the process enacting the minimisation of such prediction errors (Friston, 2005a; Friston and Kiebel, 2009c). Under this hypothesis, minimising prediction errors implies the presence of appropriate generative models producing increasingly accurate predictions of incoming sensations, or in terms of signal processing, that the input data can be successfully reconstructed using the (top-down) information stored in a model by minimising the residual error of yet-to-be accounted data.

Rao and Ballard's model thus presents a large amount of ideas later adopted in the literature to account for and possibly explain perception. It also serves as a first implementation of a mathematical model (i.e., an artificial neural network) to describe functions in the cortex as a combination of top-down and bottom-up information processing. A major limitation, however, is the focus on perception with a lack of immediate links to accounts of action and behaviour even in more recent proposals (Huang and Rao, 2011; Spratling, 2016).

2.2.3 Predictive Processing

Predictive processing (PP) is a label proposed in Clark (2013) where he suggests that

> "...rather than resting with the more common usage "predictive coding" I mean to highlight the fact that what distinguishes the target approaches is not simply the use of the data compression strategy known as predictive coding. Rather, it is the use of that strategy in the special context of hierarchical systems deploying probabilistic generative models. Such

> systems exhibit powerful forms of learning and are able flexibly to combine top-down and bottom-up flows of information within a multilayer cascade."

PP is essentially an umbrella term for approaches emphasising the presence of predictive mechanisms related to predictive coding schemes in neuroscience, and to the original algorithm in signal processing, holding however no specific assumption regarding their implementation (Rao and Ballard, Spratling, HTM, etc.). The definition given by Clark is based on probabilistic accounts of cognitive functions and, specifically, on (deep) hierarchical Bayesian models, even if hierarchies are not strictly necessary for the instantiation of predictive coding models (see the remaining chapters). In his definition of PP, Clark refers to predictive coding mainly as the signal processing compression strategy. As mentioned in Spratling (2017) however, in Rao and Ballard's model PC is already depicted as the hierarchical probabilistic model of information flow in the visual cortex, extending the original algorithm from digital signal processing to a fully probabilist formulation with precision-weighted prediction errors. One could say that Rao and Ballard's definition already captures Clark's proposal, but "predictive processing" represents nonetheless a useful label to avoid confusion with PC as an algorithm for data compression. Clark also seems to explicitly address learning as a distinctive trait of probabilistic generative models in hierarchical systems. While most certainly true in several contexts, I personally think that the word "inference" might have been more appropriate here. In my view, the most basic way of combining top-down and bottom-up flows of information stems from appropriate (Bayesian) inference processes (cf. Kalman filtering), while learning need only be invoked for more complex problems.

Clark's definition differs in some fundamental ways from the free energy principle. This principle is connected to a set of ideas (proposed to be) connected to the thermodynamics and statistical mechanics of self organisation and homeostasis for biological systems. Similar ideas are then applied to the formulation of brain theories, in particular the minimisation of variational free energy, which only under simplifying assumptions reduces to predictive coding schemes. The definition of PP lacks, in this sense, an underlying interpretation in terms of statistical mechanics and focuses instead on proposals of the brain as a Bayesian inference machine, closer to Helmholtz' hypothesis and work of analysis by synthesis. A significant contribution of Clark is however, the attempt of building a bridge between Bayesian models and 4E theories of cognition that is especially crucial in my work. In his work, he makes a distinction between a perception-centric interpretation of PP, what he defines as conservative predictive processing (CPP), and an action-oriented (Clark, 1998) reading of the same ideas, radical predictive processing (RPP) (Clark, 2015a).

Conservative predictive processing

Conservative predictive processing (CPP) is a reinterpretation of classical concepts of perception as entirely guided by external stimuli, with no direct role for behaviour. It largely corresponds to initial models of predictive coding for perception (Rao and Ballard, 1999; Lee and Mumford, 2003; Friston, 2005a) and ideas following these first attempts (Huang and Rao, 2011; Hohwy, 2013; Spratling, 2016). There is, nonetheless, a clear distinction with the more traditional purely bottom-up models of perception (Marr, 1982) since predictive coding suggests that top-down signals are crucial in defining perceptual processes. However, these signals are only produced for the sake of improving the estimate, or (perceptual) inference, of incoming signals. CPP rests on models of perception, while behaviour and motor control are often not considered. When they are, they are simply subsumed by the main goal of estimating the latent states of sensory observations, building generative models capturing the complexity of the environment. Agents in this case are described as "perception machines" (i.e., passive agents, see Chapter 1) whose only job is to capture and encode the richness and complexity of their environment. This creates a GOFAI-like reasoning system that allows an agent to simulate sophisticated cognitive tasks on an internal (generative) model that, essentially, mirrors the world. The only true novelty introduced by PP interpretations is the explicit use of top-down information flows in its generative models, inspired by predictive coding accounts of cortical activity. In this interpretation, PP is thought to be a scheme for the construction of very accurate and meticulous world models that serve higher purposes such as planning, attention and decision-making. Action is vicariously implemented based on powerful and accurate models of the environment that can be seen as detached from the world itself. The external milieu is essentially only used during the initial construction of internal models, assuming it is possible to encode all of the properties needed to plan and perform tasks whenever an agent requires it. As long as enough complexity is captured by the internal representations, interactions with the environment are not central.

Radical Predictive Processing (RPP)

Radical predictive processing (RPP) turns this passive approach around in an attempt to meet ideas closer to action-oriented mechanisms more popular in 4E theories of cognition (Engel, Friston, and Kragic, 2016). Effective generative models are the ones that allow agents to achieve their goals rather than the ones exhaustively encoding information about the world. Agents are, once again, seen as teleological systems (Rosenblueth, Wiener, and Bigelow, 1943), rather than as problem solvers (Newell, Simon, et al., 1972). "Good enough" (Loeb, 2012), parsimonious models should be preferred over more accurate and costly ones since they can make predictions that allow purposeful behaviour, which in this case is central to the general description of an agent (see the complexity-accuracy trade-off in Chapter 3 and the model in Chapter 5). Rich models are often very hard to build/learn, using a considerable amount of effort and resources, e.g., energy and time. Furthermore, in a fast-paced environment complex but slow world models may affect the coupling of agent-environment systems, while simpler but faster "heuristics" may effectively maintain their dynamics (Ashby, 1958).

In RPP, "models" may even entail some form of systematic misrepresentation (Wiese, 2016) or a severely limited objective understanding of the world dynamics as in Chapter 5, as long as these are ecologically useful in some sense, cf. Umwelt (Clark, 1998). In his work, Clark tries to connect RPP to 4E theories of cognition suggesting that RPP models have little to share with sophisticated classical interpretations of cognition having the brain in the driving seat (cf. CPP). Different levels of accuracy may affect the role of the brain, that in some cases can be secondary to teleological functions emerging purely from dynamical interactions with the environment (see Chapter 5). This may still not be sufficient to meet some of more radical interpretations of 4E cognitive science (e.g., Maturana and Varela (1980), Chemero (2011), and Newen, De Bruin, and Gallagher (2018)), where the concept of models for an agent is entirely banned, but it certainly offers an alternative way of considering agents *being* (Bayesian) models of their environment. Agents do not need to be mirrors of the world to define any useful interaction with it. On the contrary, generative models can, and in most cases should, include simple heuristics that only make sense in agent-environment coupled systems, see Chapter 5 and Chapter 6.

2.3 Action as control

In the last few decades, hypotheses inspired by control theory have emerged as dominant theories of action in neuroscience (Kawato, 1999; Wolpert and Ghahramani, 2000; Todorov, 2004; Scott, 2004; Körding and Wolpert, 2006; Körding, 2007; Franklin and Wolpert, 2011; Wolpert, Diedrichsen, and Flanagan, 2011). In this light, motor control and behaviour are seen as processes of regulating variables to a target (Powers, 1973).

The problem of regulating variables to a desired state is an ancient one, but its formalisation is relatively recent. Maxwell provided one of the first formal treatments of the problem with direct applications to the control of valves in steam engines, and in particular for the Watt-governor (that in Maxwell analysis was defined as a *moderator* rather than a *governor* since the torque was proportional to speed error rather than its integral (Maxwell, 1868)). The problem of control is common to many different research areas, from economics to biology, engineering, robotics etc. and is thus historically defined in slightly different ways. Control theory as a field is also a combination of knowledge acquired in different areas of application and therefore hard to comprehensively describe. For the purposes of this thesis, I will give only a brief overview of the fundamental phases that have defined radically different approaches to control theory (e.g., classic and optimal/stochastic) and their applications to the natural sciences. I will then describe specific examples of some modern attempts to unify information and control theory that are especially relevant for comparisons with the free energy principle in the natural sciences and introduced at the end of this chapter.

2.3.1 Classical control

Classical control is mostly concerned with linear time-invariant (LTI) systems, often in a single-input single-output (SISO) set up. Some simple applications include, for instance, thermostats (regulating the temperature of a system) or cruise controllers (regulating the velocity of a system). The focus on LTI models was mostly due to technical reasons, since at the time most of the regulation techniques were developed (roughly from 1800's to 1900-40's), computers were still not available and engineers performed their calculations on paper. Controllers were thus mostly built using heuristic methods, strongly relying on testing with post-hoc analysis of different attempts usually carried out using frequency domain techniques, checking for the stability of a solution with Laplace/Fourier transforms, pole placement and the Nyquist criterion among others (Åström and Murray, 2010). One of the core ideas introduced in classical approaches is the concept of **negative feedback**, defining control based on a signal fed back into a system to reduce fluctuations of the variable to be controlled. This idea directly connects to homeostasis in the biological science, as classically formulated by Bernard and Cannon and influenced the early cybernetics movement (Wiener, 1961; Ashby, 1957) and consequently frameworks such as perceptual control theory (Powers, 1973; Carver and Scheier, 1981), directly inspiring PP and the FEP (Seth and Tsakiris, 2018). Particularly relevant to my work is one of the most popular methodologies extensively developed at the inception of control theory in the last century: Proportional-Integral-Derivative (PID) control, see especially Chapter 6. This method (especially the integral term) was first analysed by Maxwell, who sought the necessity of the integral action for problems of steam regulation in governors (Maxwell, 1868), but was further developed and extensively applied much later on, when better design methods emerged. This controller builds on the core negative feedback component of classical control but extends it by producing an output control signal depending not only on the (negative) error between the target and the true value of a variable (here called the "proportional" term), but also on its temporal integral and derivative. The integral part essentially deals with steady state errors produced by unknown external disturbances (Sontag, 2003), while the derivative term regulates the fluctuations of the error that can arise during the control process (Åström and Murray, 2010). PID control is gaining popularity nowadays in the context of systems biology, where this mechanism has been found in different living systems as a simple heuristic for robust homeostatic regulation or "perfect adaptation" (Yi et al., 2000; Sontag, 2003; Chevalier et al., 2018) and in neuroscience/psychology as an extension of the traditional delta-rule/Rescorla-Wagner model (Ritz et al., 2018), see Chapter 6.

2.3.2 Optimal control

With the advent of the digital age, new methods became soon popular, given their mathematical formulation easily implemented on computing machines. Modern optimal control theory originated in 1950-60's with the mathematical formalisations by Pontryagin (Pontryagin et al., 1962) and Bellman (Bellman, 1957b). Connections to classical mechanics and to work by Hamilton and Jacobi among others (Todorov, 2006; Sussmann and Willems, 1997), however, place its historical roots a few centuries earlier. Optimal control defines the problem of finding a policy (i.e., a sequence of actions) for a given system or agent that optimises a criterion describing a certain goal for the system. As this simple definition already implies, optimal control deals with problems that are generally more complex than the ones tackled by classical control theory. Solutions to these problems can define in fact *policies* rather than single independent actions, dealing with time-dependent trajectories towards a goal. Actions in the classical control sense can, in this light, be seen as essentially time-independent since they are not contingent on actions taken in the future and their consequences on future observations. For example, solving a level of a Mario game requires the main character to take different actions over time that generate non-trivial outcomes in the future. Deciding on a single action at a time, only based on current world states, will soon lead our character to certain demise. In this case, one needs to consider the actions taken possibly far into the future to complete a level. A thermostat on the other hand, will simply decrease the temperature if it's too hot and increase it if it's too cold, following the same heuristic rule over time and with states far ahead in the future generally not influencing actions of the here and now. By using Bellman (for discrete time problems) or Hamilton-Jacobi-Bellman (for continuous time problems) equations, the optimal policy for a problem is determined through Bellman's optimality principle and the associated iterative Dynamic Programming formulation (Todorov, 2006), defining the so-called value or cost-to-go function (i.e., the optimal criterion to optimise). Bellman's approach is also well known for the *curse of dimensionality* since the Dynamic Programming method requires computing the objective function for each combination of values, making the problem practically infeasible in highly dimensional spaces (Bellman, 1957b). Pontryagin principle, on the other hand, is not afflicted by the same problem and can be seen as a generalisation of variational methods from classical mechanics (thus also the the connections of optimal control to work by Hamilton and Jacobi (Todorov, 2006)). Its generalisation to real-world problems with stochastic components, unlike Bellman's formulation, remains however more problematic (Todorov, 2006).

Following the especially popular optimal control view in the natural sciences, organisms are often depicted as minimising a cost function (i.e., an optimality criterion) over time, representing a measure of performance for the achievement of a certain goal. For instance, smooth reaching hand movements can be described by the minimisation of a cost function based on the rate of change of acceleration, or jerk, of hand movements (Franklin and Wolpert, 2011). One of the main limitations of this approach is however the lack of tools to deal with noise and uncertainty, inherent properties of real-world problems.

2.3.3 Stochastic optimal control

In his pioneering work, Bellman also defined the basis for the formulation of the stochastic version of the optimal control problem (Bellman, 1957a), starting what is now known as stochastic optimal control (Åström, 1970; Anderson and Moore, 1990; Stengel, 1994; Todorov, 2006; Kappen, 2011). This formulation emphasises the fundamental probabilistic nature of optimal control/decision making problems, with outcomes that are (at least partially) stochastic and not under a system's direct control. The dynamic programming formulation can still be applied to this case, but as pointed out previously, Pontryagin principle is mostly limited to deterministic problems. As previously mentioned, the field of control theory includes the contribution of different research areas and in particular for stochastic optimal control, of estimation and information theory due to the probabilistic formulation of the problem. Work by Kalman (Kalman, 1960c) vastly contributed to the field, with the definition of properties such as observability and controllability of a system, the former expressing the degree to which states can be estimated from noisy observations, and the second one representing the degree of control of a system given when different manipulations are applied. Kalman first noticed also that his newly defined filter was the dual problem of a well known example in linear optimal control theory: the linear quadratic regulator (LQR) (Kalman, 1960b), with both problems based on the solution of Riccati equations. In Chapter 7 I will discuss some of the possible implications of this finding for the cognitive sciences. In particular, the definition of the linear quadratic Gaussian (LQG) method and the so-called *separation principle* of control theory, central to applications of (stochastic) optimal control to neuroscience (Todorov and Jordan, 2002; Todorov, 2004; Scott, 2004), will have non-trivial implications for some principles of 4E proposals. The free energy principle itself strongly relies on this duality and brings it to the realms of neuroscience and cognitive science with a set of hypotheses generated from empirical studies (see, again, Chapter 7).

In (stochastic) optimal control, assumptions on the nature of the problem at hand lead to the use of different cost functions for the construction of different control policies (i.e., sequences of actions) towards a goal. One of the most fundamental distinctions with repercussions on the natural sciences is the one between classes of cost functions in control theory (Todorov, 2004; Ahissar and Assa, 2016; Buckley and Toyoizumi, 2018). This distinction is based on the absence/presence of real-time feedback from the environment, defining open and closed-loop control respectively, see (Åström and Murray, 2010) for an introduction. Open-loop control methods rely on complex internal models that accurately mimic the external dynamics including the effects of feedback, and allow for pure internal feedforward planning. In this type of control models, perceptual processes can be ascribed to the transduction of sensory input into some internal (e.g., neural) representation and are often depicted as estimators or forward models-estimators pairs (see section 2.2). These representations produce then (motor) actions via a fixed, pre-programmed, controller (Todorov, 2004) or inverse model (Jordan, 1996; Kawato, 1999) that does not take into account feedback from the environment. The sequential nature of estimation-modelling (planning)-control is consistent with the traditional classical sandwich of cognitive science, where estimation and control are processes encapsulated into separate modules used in sequential order, and the effects of feedback from the environment are though to be deterministic, slow and stable enough to be successfully modelled internally. Closed-loop control is based on the same modules used in the open-loop case (i.e., estimators and controllers) but includes, in contrast, fast-paced feedback from the environment. This allows such framework to elegantly tackle different sources of noise, internal fluctuations and uncertainties (Todorov and Jordan, 2002; Todorov, 2004; Åström and Murray, 2010; Pezzulo, Rigoli, and Friston, 2015; Ahissar and Assa, 2016; Buckley and Toyoizumi, 2018) although its performances in presence of delays are still under scrutiny (Jordan, 1996; Franklin and Wolpert, 2011). Closed-loop control appears to be closer to 4E views because of the presence of a real-time feedback mechanism that highlights the crucial dynamic interactions of a system with its environment and the insufficiency of purely feedforward models. We however will argue that the most common implementations of closed-loop optimal control are still more consistent with the traditional, sequential, view of action and perception mainly due to ideas based on the *separation principle* of control theory (Stengel, 1994), introduced and discussed in Chapter 7. In other research areas such as artificial life, adhering more closely to 4E ideas, different models provide examples where this separation is not assumed, not present or at least not easily detectable/found (Beer, 2003; Iizuka and Ikegami, 2004; Harvey et al., 2005; Di Paolo and Iizuka, 2008).

2.3.4 Reinforcement learning

Reinforcement learning (RL) provides a complementary perspective to optimal control methods, bringing together a long tradition of studies in psychology and animal learning including work by Pavlov, Thorndike and Skinner among others (Pavlonian conditioning, the law of effect and operant conditioning) and the optimal control formulations summarised above (Sutton and Barto, 1998). In their simplest form, models of reinforcement learning can be seen as defined for the same problem tackled by (stochastic) optimal control: to determine the optimal policy for a system having some goal. A major difference between RL and (stochastic) optimal control lies in the fact that the latter does not deal with the problem of learning: it is generally assumed that all parameters of the model of the system to regulate are available (see however Kappen (2011) for a more generous definition of learning, describing what I define here as inference+learning). Optimal control heavily relies on the formulation of *dual control* problems in the context of (partially) unknown processes by Feldbaum in the 1960's, defining the exploration-exploitation dilemma for control theory (Wittenmark, 1995). Since no analytical, general solution to this problem exists, optimal control methods use simple heuristics such as the iteration of estimation and control phases to generate approximate but well understood solutions (Kappen, 2011). Reinforcement learning, on the other hand, adopts in general a position closer to *adaptive* control (Kappen, 2011), looking for a solution by trial-and-error, learning and inferring properties of a system by repeating the same task as if time and number of trials were not a constraint (Sutton and Barto, 1998) and unlike dual control, where knowledge about a system must be acquired while attempting to regulate it. The approach adopted by most optimal control methods focuses on the definition of analytical solutions to the control problem, having thus limited applications since this mathematical tractability often imposes strong constraints, e.g., linearity. Standard RL, on the other hand, is less concerned with analytical solutions and provides a more general set of tools for dealing with real-world problems where conditions of observability and controllability (Kalman, 1960a) cannot be verified and agents simply have to find the best (even if heuristic) policy. In particular, if full observability of a system is not available or if some of the parameters of the system to control are unknown, RL provides a set of methodologies that can be used in such cases, from policy-iteration or value-iteration, to Q-learning, SARSA, etc. (Kaelbling, Littman, and Moore, 1996; Sutton and Barto, 1998). These approaches belong to the class of methods defined as *model-free* reinforcement learning, in contrast to a series of other methods falling nowadays under the name of *model-based* (Sutton and Barto, 1998). While the details of the different proposals are not especially relevant for the remainder of this thesis, this distinction is important as it resonates with a similar difference in approaches to the definition of the free energy principle, denominated "agency-free" and "agency-based" respectively (Friston, Samothrakis, and Montague, 2012a) or "belief-free" and "belief-based" in (Friston et al., 2016a). Model-free approaches in reinforcement learning refer to the set of methodologies where agents simply accumulate information regarding valuable states by trial-and-error and use these heuristic knowledge of valuable (explored) states to construct policies. Model-based methods, on the other hand, attempt to build (i.e., learn) a model of the state transition probabilities and of the value function to determine policies with a candidate model of the world dynamics that can also work for unexplored states (Kaelbling, Littman, and Moore, 1996). Model-free methods can be associated to agency-free formulations of the free energy principle and the examples provided in this thesis can be seen in this context. In the models implemented here however, there is no explicit trial-and-error phase, with generative models either already encoding the relevant properties about the policy as priors provided by the experimenter, or representing trivial time-independent policies (e.g., a thermostat) as discussed in section 1.1.2. An example of the former can be found, for instance, in Friston, Daunizeau, and Kiebel (2009) (see also the discussion of this model later in the chapter), while the latter can be seen essentially in all of the implementations in the following chapters. Model-based methods are more aligned with an extension of the free energy framework, incorporating an account of expected variational free energy minimisation on future and yet-to-be seen observations, and explicit distributions over (hidden) control states in the future with state transition probabilities that can be inferred for more complex problems (Friston et al., 2015; Friston et al., 2016a; Friston et al., 2017). As explained in section 1.1.2 however, this is not be treated in this thesis.

2.3.5 Other relevant approaches to control

Several other approaches to problems of control have also been proposed in the literature. In this section I will present a brief overview of only but a few examples using information theoretical definitions for control, including specifically: KL-control, empowerment, homeokinesis and predictive information.

KL-control KL-control (Kappen, Gómez, and Opper, 2012) defines a class of control problems formulated as the minimisation of a Kullback-Leibler divergence (Kullback and Leibler, 1951; Bishop, 2006). This minimisation is aimed at decreasing the deviation of a probability distribution p representing the *controlled* dynamics of a system, i.e., the system regulated by a specified controller, from a distribution q representing the *uncontrolled* dynamics of a system, i.e., a distribution of the system in its desired state with no actions/forces (i.e., , control) applied. The body of work covered by KL-control, see also Todorov (2008), Todorov (2009b), and Todorov (2009a), is largely overlapping with the more general definition of expected free energy (Friston et al., 2015) and time-dependent policies, thought to generalise KL-control among other approaches (Friston et al., 2017). This method lies, therefore, outside the scope of what I cover, as discussed in section 1.1.2. Furthermore, while KL-control provides a better framework for time-dependent policies, these approaches heavily rely on discrete time systems and thus are crucially different from the ones used in this thesis.

Empowerment Empowerment is a measure defining the ability, or "power", of an agent to knowingly change its environment (Salge, Glackin, and Polani, 2014). It is defined as the Shannon channel capacity of, i.e., the maximisation of mutual information between, an agent's actuators states and its sensations in the future (i.e., the consequences of its actions). It was initially described for discrete-time systems (Klyubin, Polani, and Nehaniv, 2005a; Klyubin, Polani, and Nehaniv, 2005b) and later extended for the continuous-time case (Jung, Polani, and Stone, 2011). One of the main features of this approach is the definition of a candidate intrinsic motivation for biological agents: to maximise empowerment. The possible implications cover hypotheses in evolutionary theory, behavioural studies and AI implementations (Salge, Glackin, and Polani, 2014). This constitutes an advantage over the formulation of the free energy principle used in this thesis, since the minimisation of surprisal alone generates time-independent policies that do not, effectively, take into account the effects of actions on future states as discussed in section 1.1.2 and Kappen, Gómez, and Opper (2012) and Friston, Samothrakis, and Montague (2012a). On the other hand, the framework based on expected free energy extension can be compared to empowerment (Biehl et al., 2018), with the two theories providing different candidate intrinsic motivations: maximisation of empowerment vs. minimisation of *expected* free energy. This comparison remains also outside the scope of my work. Furthermore, the computation of empowerment is known to be particularly complex and hardly scalable due to the combinatorial complexity of states in grid worlds in discrete space and time proposals and, in the continuous-time formulation, due to Monte Carlo approximations (Jung, Polani, and Stone, 2011).

Homeokinesis Homeokinesis is a generalisation of the classic notion of homeostasis, where the target is not *stasis*, i.e., , a fixed or steady state, but rather *kinesis*, i.e., , an equilibrium kinematic regime of a system and its constituents with the environment (Der, Steinmetz, and Pasemann, 1999; Der and Martius, 2012). This approach defines adaptive behaviour via the minimisation of an error function representing how well a system can predict the consequences of its own actions. To avoid trivial solutions, "do nothing" (cf. the dark-room problem in Chapter 4), the error function includes also a drive to maximise the sensitivity of an agent's sensory system, such that its actions can "destabilise" the world to introduce novel observations. In robotics and artificial life, the notion of homeokinesis has been proposed as a principle for the emergence of (intrinsically motivated) complex behaviour in artificial, and possibly biological agents (Der and Martius, 2012). Homeokinesis introduces a notion of prediction error minimisation similar to the one adopted by the free energy principle and can be seen as closely related to a dynamical systems interpretation of the FEP (see Chapter 6). However, until the introduction of predictive information (see below), this framework lacked the flexibility afforded by the mathematical formulations adopted by the FEP (e.g., information/probability theory).

Predictive information Predictive information is a measure of the correlation between past and future of a variable, more specifically their mutual information (Bialek, Nemenman, and Tishby, 2001). In the context of adaptive systems, it is proposed to quantify the total information of past experience that can be used to predict future events, i.e., , how much past observations help in the prediction of future ones. The intuition of predictability of future observations behind predictive information is similar to the one introduced by homeokinesis, and the two frameworks have been shown to be formally connected Ay et al., 2008. As in the case of empowerment, predictive information has also been proposed as a possible intrinsic motivation (Ay et al., 2012; Biehl et al., 2018). Similarly to empowerment, predictive information offers a broader perspective on adaptive behaviour when compared to the formulation of the free energy principle adopted here, which in itself offers no effective implementation of intrinsic motivation mechanisms (see also discussion on dark room in Little and Sommer (2013)), defining essentially just “reflex-based” or “agency-free” agents (Friston, Samothrakis, and Montague, 2012a). As with the previously discussed measures, a better comparison could be introduced with the *expected* free energy formulation (Friston et al., 2015; Friston et al., 2017) but once again, this will not be covered in this thesis. See Biehl et al. (2018) for discussion on these ideas.

2.4 The Free Energy Principle (FEP)

Initially proposed by Friston, Kilner, and Harrison (2006) and further discussed in a series of papers including (Friston, 2009; Friston, 2010b; Friston, 2012; Friston et al., 2015; Friston et al., 2016a; Friston et al., 2017), the free energy principle (FEP) is proposed as a unifying brain theory with roots in thermodynamics and statistical mechanics. According to the FEP, and following Schrödinger classical idea of negentropy (Schrödinger, 1944), biological systems only exist far from thermodynamic equilibrium and in some way manage to self-organise and maintain a certain level of order through homeostatic regulation. The way this is portrayed in Friston’s work implies that such systems try to minimise the Shannon (information) entropy of their sensory input which would keep thermodynamic entropy within boundaries (Sengupta, Stemmler, and Friston, 2013; De Ridder, Vanneste, and Freeman, 2014). The brain is thought to be one such system and the information entropy is in this case defined as the uncertainty of sensory input (Friston, Kilner, and Harrison, 2006).

In this thesis, I will only tackle the implications of the FEP in the realm of cognitive (neuro)science, trying to place them in the context of ideas reviewed so far in this chapter. Some of the hypotheses formulated under the FEP, for instance simulations of self-organising systems (Friston, 2013; Friston et al., 2015), will not be part of my presentation. A full mathematical derivation of the FEP following Buckley et al. (2017) and discussing some extra matters related to Friston, Trujillo-Barreto, and Daunizeau (2008), Friston (2008b), Friston (2008a), and Friston et al. (2010b) is then presented in the next chapter, as a methodological background for the presentation of my results further in this thesis.

As I will show in Chapter 3, under simplifying assumptions, mainly the Laplace encoding (Buckley et al., 2017), the FEP reduces to predictive coding as proposed by Rao and Ballard (Rao and Ballard, 1999) and discussed earlier in this chapter (see also Friston and Kiebel (2009c)). One of the main contributions of the FEP is then the extension of predictive coding models to accounts of action. The introduction of “active inference" (Friston, 2009; Friston et al., 2010a) represents in fact a way to unify perception and action in a cohesive mathematical framework where differences between the two of them almost vanish.

2.4.1 Active inference

If perceptual inference implies that models can be updated to better infer the world, active inference (Friston, 2009; Friston et al., 2010a) suggests also that acting on the environment may change it so to be better described by a given model. Active inference, in fact, introduces a second way in which prediction error, or variational free energy, could be minimised. In more traditional views of predictive coding, perception is the only process contributing to the minimisation of prediction errors. On this view, an agent can only update predictions of its generative model to explain the environmental causes of those errors and in doing so, the errors will be explained away. Active inference introduces, in this context, a way for agents to change the signals representing sensory input to better fit their predictions. Agents thus actively interact with the environment to produce sensations that generative models can predict. Behaviour is generated through interactions with the world that are defined in terms that are consistent with the perceptual account of the FEP. Motor commands are explained as predictions generated by the same generative model at a proprioceptive level compared with actual proprioceptive input (Friston, 2011; Adams, Shipp, and Friston, 2013).

2.4.2 Active inference agents

The free energy principle is a relatively recent approach proposed for the study of cognition and as such, only a relatively small number of models, especially agent-based ones, have been implemented for the investigation of the claims made in its formulation. As a proposal for a general theory of brain functioning, most of the models developed under the FEP or that can be considered aligned with it, are based on implementations of predictive coding models of cortical functions, for example in the visual (Rao and Ballard, 1999; Lee and Mumford, 2003; Spratling, 2008), auditory (Baldeweg, 2006) or the motor cortices (Shipp, Adams, and Friston, 2013; Adams, Shipp, and Friston, 2013). These models are largely focused on explanations and hypotheses regarding information processing in the cortex and as such, they emphasise possible implications for perception and are mostly concerned with building accurate (although mostly functional) descriptions of cortical activity. In line with these models and the FEP/active inference proposal more in general, the emerging field of computational psychiatry also attempts to follow a similar route, in this case for the mechanistic investigation of psychiatric disorders (Montague et al., 2012; Stephan and Mathys, 2014). My focus in this project is however on building synthetic agents and thus will be radically different, although working within the same mathematical framework. Some of the examples of agent-based models based on the FEP are here reviewed, in particular solutions to the mountain car problem proposed by Friston and two implementations of novel systems, the linebot and the infotropic machine.

The mountain car problem

In a series of papers (Friston, Daunizeau, and Kiebel, 2009; Friston, 2010b; Friston et al., 2010a; Friston and Ao, 2012; Friston, Adams, and Montague, 2012; Friston, Samothrakis, and Montague, 2012a), active inference is proposed for the solution of a classical control problem in reinforcement learning, commonly referred to as “mountain car" or "park on a hill" problem (Sutton and Barto, 1998). The agent is in this case a car that needs to drive up one of the two sides of a steep valley after having been placed at the bottom valley. Its actuators are however under-powered, so that even at full throttle the car cannot get out on either side. By going full throttle in one direction first, the car can instead gain enough momentum and swing up to the opposite side. The solution is considered non-trivial since the agent needs to first get away from the target location to be able to reach it later on.

Most of the implementations proposed by Friston focus on the definition of active inference solutions to this task, often redefining some of the several different methods used in reinforcement learning to solve the problem. This redefinition entails, essentially, the use of priors on a target location (cf. KL-control, discussed earlier) instead of using value functions as in classic model-free reinforcement learning (Friston, Daunizeau, and Kiebel, 2009). It is claimed that this approach generalises classical model-free reinforcement learning algorithms and optimal control formulations (Friston and Ao, 2012) by replacing value functions with ad-hoc priors and thus avoiding backward induction typical of path integral approaches defining value as a function of future states. It is also argued that “itinerant" as opposed to “fixed" policies, the latter typical of reinforcement learning set ups, can be implemented in active inference thanks to the formulation of state-space models in generalised coordinates of motion, see Chapter 3. These policies generalise fixed target points by implementing, essentially, “target trajectories” (see Chapter 6) and allow for the definition of chaotic behaviour and exploration/foraging similar to what is found in insects (Friston, Adams, and Montague, 2012). While interesting for its generalisations of fixed target states, this formulation is deeply problematic since it assumes pre-existing knowledge of appropriate policies that can be effectively implemented by an agent (Friston, Daunizeau, and Kiebel, 2009; Friston and Ao, 2012). Friston often argues that these policies emerge on longer (evolutionary) time-scales (Friston, 2010a; Friston and Ao, 2012) but, practically, this nativist proposal doesn’t explain how they came about if not in terms of tautological and self-referential arguments of the kind: if they didn’t, such systems wouldn’t be there to begin with. Moreover, most of these implementations assume a very deep isomorphism between the physical laws of the (virtual) world defined as *generative process*, and the agent’s encoding of such mechanisms, i.e., the *generative model*. From a detailed analysis of all the cited papers implementing solutions to the mountain car problem under active inference, I then propose to group solutions based on three different approaches for which I highlight features and limitations.

Ad-hoc environments In Friston, Daunizeau, and Kiebel (2009) and Friston et al. (2010a), the agent is placed in a ad-hoc environment built specifically for training and explicitly called “controlled environment". This simulated environment implements a generative process, the set of physical laws describing the environment, different from the ones adopted in the original problem formulation. These laws are found independently by the experimenter (Friston, Daunizeau, and Kiebel, 2009) via a minimisation of a KL divergence between desired and controller environments, as in KL-control (Kappen, Gómez, and Opper, 2012). Effectively, this method solves the task independently and outside of the simulation, and is used for the purpose of an initial phase of (supervised) training. Once the expected behaviour is learnt in this fictitious environment and encoded by the agent in terms of priors, the same agent is then placed back into the original simulated world where it acts based on these now fixed (and strong) priors. It remains unclear how the problem could be solved without such “controlled environment”, as defined by the authors. Using this approach, in fact, the authors effectively built the optimal environment for training which clearly takes away much of the burden from the agent as admittedly stated in Friston, Daunizeau, and Kiebel (2009): “...it could be said that we have done all the hard work in creating a controlled environment; in the sense that this specifies an optimum policy, given a desired equilibrium density (i.e., , value function of states). This may be true but the key point here is that the agent does not need to optimise a policy...".

Complex itinerant policies In Friston (2010b) and Friston and Ao (2012), the authors show an example of itinerant policies in generalised coordinates of motion. In Friston (2010b) an agent is provided with a prior that essentially encodes a belief in a world with “negative friction” away from its target outside the valley. The car thus explores the environment in an attempt to constantly increase its acceleration and velocity due to its beliefs about the effects of a (fictitious) negative friction. To do so, it will start moving from side to side: the only way to implement beliefs about a force essentially pushing it around, and eventually reach the target position where the prior encodes instead a strong friction in order to stop the car. A smart initialisation of the position of the car halfway up the hill ensures then the initial movement and a prompt solution to the problem. In Friston and Ao (2012), the authors implement a time-dependent constraint in analogy to “satiety”, simulating goals that are interesting for the agent only for a limited time and thus promoting exploration, solving the mountain car problem indirectly. The agent wants to get out of the valley but gets stuck on one side, satiety decreases after a short time and changes the prior/desire to make the opposite side more interesting and allowing the car to gain momentum by going full throttle into the opposite direction. Satiety then increases, making all states equally attractive and thus forcing the agent to explore as many of them as possible, which will quickly make satiety go down once again, etc.. A new cycle will thus start, with the car attracted by an itinerant fixed point that periodically appears/disappears. As pointed out previously, it remains unclear to imagine how both these complex prior beliefs, negative friction and satiety, would emerge from simple interactions of an agent with the environment. More experiments could shed light on this idea, prescribing some mechanisms of learning with almost no initial knowledge available as in standard model-free reinforcement learning (Sutton and Barto, 1998). In Friston, Samothrakis, and Montague (2012a) the authors attempt to bypass this problem by assuming that learning *must be supervised* with ad-hoc pre-built environments playing the role of teachers and perfect representations of the world learnt by an agent, as previously done in Friston, Daunizeau, and Kiebel (2009) and Friston et al. (2010a). The authors state: "One might ask where these worlds come from. The answer is that they are created by teachers, parents and conspecifics. In robotics and engineering, the equivalent learning requires the agent to be shown how to perform a task" (Friston, Samothrakis, and Montague, 2012a). However this can only account for supervised learning, imitation, which represents but one possible way agents can learn about their milieu. This perspective is shared in various areas of machine learning and robotics but it certainly does not help defining mechanisms by which systems learn to actively interact with the environment without explicit, external supervision.

POMDP implementations In Friston, Adams, and Montague (2012) and Friston, Samothrakis, and Montague (2012a), the focus shifts towards discrete implementations of the problem showing the formal equivalence between optimal control in Partially Observable Markov Decision Problems (POMDP) and active inference processes. In these papers, the authors show that given a target location (outside the valley) and a set of transition/output matrices implementing how states change over time with different actions and how these related to observations, the optimisation process can be cast as Bayesian inference. However, once again the agent is equipped with an accurate generative model with transition/output matrices mirroring the actual description of the mountain car problem. The presence of known transition/output matrices allows to bypass the problem of learning complicated priors replacing value functions and defining a model over which policies are determined. In continuous-times set ups, this formulation remains ill-defined and it is not very clear how the same process could be implemented in a computationally affordable way. As some of the same authors suggest, this formulation may be complementary to the continuous one adopted in this thesis (Parr and Friston, 2018c), proposing that decision making problems are, perhaps, intrinsically discrete in nature and only peripheral functions/reflexive behaviour need to be modelled as continuous.

The linebot

The “linebot" is an agent-based model inspired by the FEP proposed in McGregor, Baltieri, and Buckley (2015). An agent moves in a one dimensional toroidal world with a source of chemical whose concentration decays away from the position it is placed. The agent is provided with one sensor that provides a single bit of information, namely high or low reading of the chemical with a probability proportional to its concentration. The agent is implemented with a desire system that represents its preference to settle down at a specified position with an appropriate concentration for its own survival. The task is made hard by the fact that the agent starts with no information about its own location, information that can only be inferred by moving and sensing different levels of the chemical. On top of that, the line world presented in McGregor, Baltieri, and Buckley (2015) implements a symmetrical spatial decay, meaning that pairs of different positions will give the same probabilistic encoding of the sensory reading. The space is discretised in order to allow a simplified but well-defined behaviour by this agent, moving in its environment by fixed spatial units. Within the same unit, the chemical level is kept constant.

Following this set up, a series of simulations were presented mainly for didactic purposes in order to investigate different aspects of the FEP. The example is also simple enough to allow exact Bayesian inference, compared in one of the simulations to the variational approximation used by the FEP. Some other approximations are, however, not considered (mainly the Laplace assumption, see Chapter 3). The authors, in this case, proceed with the minimisation of variational free energy by exhaustive computation of probabilities of all possible states and actions, making scalability an issue (simulations on a 3-by-3 two dimensional grid took already a few hours). While extremely useful for initial exploratory research, in this set up there is thus little room for the implementation of more complex behaviours. In the continuous case, to make the optimisation tractable, assumptions such as the Laplace encoding are usually included, together with gradient descent methods for the efficient optimisation of the free energy functional (Buckley et al., 2017). In the discrete case (extending the formulation to *expected* free energy), the equations for the minimisation of free energy are either determined analytically (Friston et al., 2015) or follow the same idea of gradient descents (Friston et al., 2017). Furthermore, the linebot in McGregor, Baltieri, and Buckley (2015) implements a generative model faithfully encoding the generative process/laws of the environment. A different example comes from my MSc dissertation (Baltieri, 2014), where I made explicit assumptions about an agent having limited resources, building a factorised hierarchical generative model of the environment, showing that even with simpler encodings this agent can still find its position in the environment. My work at the time did not, however, implement behaviour through active inference, having just some simple ad-hoc rule for moving towards the target, rendering those simulations not comparable to results presented in this thesis.

The infotropic machine

The infotropic machine is a model proposed by Thornton (2016). “Infotropic" is the definition given to characterise information seeking systems (Thornton, 2014). The author’s approach is rooted in information theory and provides a different set of initial assumptions with respect to other PP/FEP agent systems. In this work, Bayesian computations are not explicitly involved, although one could argue that some of the agent’s functions can be interpreted under a Bayesian framework anyway. As an example implementation, the author provides an implementation of Braitenberg vehicle 3a “permanent love" (Braitenberg, 1986), described as a hierarchical prediction machine. This vehicle approaches the source of the stimulus (light or chemical) while slowing down, stopping at the source, as in the classical formulation. The infotropic version of this vehicle implements a controller that essentially predicts the permanent love behaviour through the application of hierarchical message passing in a network of nodes computing a newly defined information measure: information payoff. This quantity is proposed as an alternative to the KL divergence typically used in variational inference schemes (see Chapter 3) to measure the informational value of a prediction, and its maximisation as an alternative to the minimisation of variational free energy. As in the case of the KL divergence, information payoff is also not a metric, since it is defined on the interval $[-1, 1]$, but unlike the KL measure, well explained in information geometrical terms for Riemann manifolds with a Fisher metric (Amari, 2016), it remains unclear how information payoff would behave, in general, over a probability space.

As stated by the author, some of the components presented in other PP/FEP frameworks are missing, namely: error units, feedforward error signals and precisions. The author then raises a few important questions regarding the necessity of such components. For instance even if different sub-populations of neurons have been empirically confirmed, the architecture describing error units as physically separate from prediction units (c.f. deep and superficial pyramidal cells in Friston (2008a)) remains largely an hypothesis. It is also unclear what feedforward and feedback signals exactly convey (Spratling, 2013; Spratling, 2017), with the more popular implementation by Rao and Ballard representing at most a functional model of cortical processing (Rao and Ballard, 1999), over which some hypotheses on neurophysiological implementations have been formulated, see for instance Bastos et al. (2012) and Keller and Mrsic-Flogel (2018). On the other hand, the dismissal of precisions is something that I consider more problematic. Precisions, in the FEP framework, represent a weighting mechanism for prediction errors that assigns a level of confidence based on estimates of external noise and internal fluctuations, and priors/desires. Thornton (2016) suggests to combine precisions with the very quantity they should provide a weight for, in this case not prediction errors but informational value units. He specifically argues that: “A prediction error calculated with zero confidence translates into an informational value of zero bits, for example." (Thornton, 2016). He then proposes that his definition of information payoff captures both concepts into a single measure. However this seems to miss out on a few important matters. For instance, the fact that some non-zero prediction errors may simply be non-zero and that only their weights vary over time (Brown et al., 2013), introducing thus more degrees of freedom. More in general, generative models can be multivariate and can present correlations between different variables than are well captured by precisions (inverse COvariances matrices). The author's interpretation may be relevant in a univariate case or in multivariate scenarios with uncorrelated variables, but the argument needs further investigation.

Other models

Friston, Samothrakis, and Montague (2012a) and Friston et al. (2015) introduced an extension of the FEP that includes beliefs about control states, fictitious actions, that are inferred by minimising *expected* free energy in the future. Unlike the formulation used in this project, in this more general framework activities like planning can emerge while inferring states of the world. In this set up, the agent's beliefs get refined with more incoming information (reducing epistemic uncertainty), allowing then the optimisation of policies to afford the minimisation of expected free energy about the future. Thanks to this approach, the authors suggest that they can formulate planning and the exploration/exploitation trade-off as emerging from the minimisation of expected free energy. It is also claimed that this trade-off can be "solved" ("Formally speaking, we resolve the exploration-exploitation dilemma by endowing agents with prior beliefs that they will minimize the expected free energy of future outcomes." Friston et al. (2015)) but effectively, agents are simply provided with appropriate priors under which this dilemma is already "solved". Over the last few years, several agent-based models have been implemented under this extended formulation. For example, Friston et al. (2015) show a virtual rat moving in a T-shaped maze. The rat is initially placed between the two upper arms of the maze. To find a reward in one of the arms, the rat needs to move back to the lower end of T-maze and acquire information about the reward's position, placed in one of the other two arms. The work presented in this paper is very promising since it tries to explain planning in terms of inference via the minimisation of expected free energy, extending work such as Attias (2003) and Botvinick and Toussaint (2012). The generative model, and especially the priors, are still largely provided but the authors suggest ways of dealing with the more general problem of learning by including an intrinsic drive to explore to solve epistemic uncertainty. The extension of this framework in terms of expectations of variational free energy, as pointed out in the Introduction, lies outside the scope of this thesis but constitutes, nonetheless, an interesting venue for future research. Some interesting directions include, for instance, an integration of the discrete expected free energy formulation and the continuous one (Parr and Friston, 2018c) and its meaning for theories of intrinsic motivation (Biehl et al., 2018).

Friston's group proposed also other models that are sometimes referred as agents but that effectively lack some of the commonly features of agent-based models, especially the presence of a clear and distinctive "environment" as separate from an entity defined as "agent" (Barandiaran, Di Paolo, and Rohde, 2009). Examples of this kind include work on Lorenz attractors, one of simplest models of chaotic systems (Friston et al., 2010a; Friston and Ao, 2012). A hierarchy of Lorenz attractors is also then implemented to model the inference of hidden properties in birdsongs (Kiebel, Daunizeau, and Friston, 2008; Friston and Kiebel, 2009a). An interesting point for this thesis is nonetheless found in Friston, Daunizeau, and Kiebel (2009), where the authors built an "agent" whose generative model is represented by a Lorenz attractor, interacting with a "world" described by dynamics that simply settle to equilibrium over time, i.e., no chaotic behaviour. Strong prior beliefs by the agent attempting to impose chaotic behaviour on the environment through the injection of input (i.e., action) in the dynamical system governing the dynamics of the world, allow this "world" to behave like a Lorenz attractor itself. While stretching the definition of agent-environment systems, this implementation represents one of the few examples (the other one found in Schwartenbeck et al. (2015)) where active inference is specifically applied in the context of coupled dynamical systems (a generative model and a generative process, "agent" and "environment") that are different in some nontrivial way, with one of them effectively driving the dynamics of the other (see also Chapter 5).

Some other implementations (Friston, 2013; Friston et al., 2015), again led by Friston's group, focus on hypotheses of self-organisation formulated under the FEP and give examples of how protocells might have formed out of very simple physical and chemical interactions in ways consistent with the FEP. Since this part of the theory is not discussed here, the "agents" implemented in this line of research will also not be examined.

2.5 Conclusion

Embodied, enactive, embedded and extended (4E) theories of the mind highlight the importance of sensorimotor loops, the presence of feedback from the environment, the distributed nature of cognition and a lack of explicit symbolic manipulations in agents. In this chapter I gave an overview of some of these ideas in light of cybernetics proposals for the study teleological systems as if they were controllers/regulators, emerging for example in Rosenblueth, Wiener, and Bigelow (1943) and Conant and Ashby (1970), that inspired some core ideas of 4E theories. I also discussed then some potential limitations of the explicit connection I make between 4E and cybernetics, operationalising my definition of model for an agent. Some areas of 4E theories seem to reject ideas deeply rooted in (optimal) control and especially work on teleology (Di Paolo, 2005; Froese and Ziemke, 2009) as addressed in Rosenblueth, Wiener, and Bigelow (1943), thought to be problematic because of the presence of extrinsic ideas such as *value*. In the formulation of the free energy principle however, value is replaced by the minimisation of sensory surprisal proposed to implement intrinsic drives (in the form of priors) for different living systems (Friston, Kilner, and Harrison, 2006; Friston, 2010b; Friston, 2012).

Using this point to support investigations of the FEP within 4E views in the next chapters, I then extensively reviewed some of the most prominent modern theories of action and perception, placing them in the context of control and estimation theory. In particular, from the inception of optimal control theory in the last century, control theorists and engineers have sought to study problems of regulation using a combination of estimation or *inference* and *control* (Åström, 1970; Stengel, 1994). The former is proposed to identify and recover features of systems to be regulated in the absence of direct observations of relevant variables, the latter to build effective regulators once enough information on the system to control is available. In the cognitive sciences, similar approaches are nowadays proposed with the goal to describe and potentially explain processes of perception *as inference* and action *as (optimal) control*. The core of this chapter reviewed these proposals in detail and set the ground for a conceptual introduction of the free energy principle and active inference. The free energy principle was thus briefly explained, highlighting in particular its attempt to unify theories of perception as inference and action as control in the cognitive sciences where the two lines of research are still fundamentally disjoint. In the remainder of the chapter, I then considered in detail several agent-based models proposed in the literature under the free energy principle and that can be compared to the models that I will introduce later in the thesis, especially in my attempt to make explicit connections to 4E theories. A full presentation of the FEP must include the necessary mathematical tools and is thus provided in the next chapter.

Chapter 3

Methods

3.1 A mathematical formulation of the free energy principle

The free energy principle rests on the idea that all biological systems must avoid dispersion in order to survive. In thermodynamics, entropy represents the degree of energy dispersal of a system, and following the second law, the entropy of a thermodynamically closed system cannot, on average, decrease over time. According to the FEP, living systems can only exist by resisting the effects of, i.e., (locally) minimising, entropy (Friston, 2012; Friston, 2013; Ramstead, Badcock, and Friston, 2017), an idea inspired by Shröedinger's inception of negentropy production to describe life (Schrödinger, 1944) and previous ideas by Clausius, Kelvin and Boltzmann among others (Boltzmann, 1974). The argument initially built on the second law is then better framed in terms of the fluctuation theorems for non-equilibrium thermodynamics: for open, dissipative systems (cf. closed ones for the second law), entropy can *locally* decrease with a certain probability, see for instance (Evans and Searles, 2002). Since direct measures of entropy in a thermodynamic sense are not accessible by an agent, the FEP proposes in its place the minimisation of information or Shannon (differential) entropy of the states ψ perceived by an agent (De Ridder, Vanneste, and Freeman, 2014), defined as:

$$H(\psi) = -\int_{\psi} p(\psi) \ln p(\psi) \, \mathrm{d}\psi \tag{3.1}$$

By decreasing the number of states it observes and thus the Shannon entropy of its observations, it is claimed that an agent also reduces the number of the physical states it can physically occupy, i.e., the agent minimises its own thermodynamic entropy by staying in a limited amount of states where only a limited amount of observations are available (Sengupta, Stemmler, and Friston, 2013)[1]. The information entropy of observations ψ itself is also not directly accessible by an agent, since it requires an integration over the ensemble of all the possible observed variables. With an ergodic assumption (Friston, 2012), equating the integration over an ensemble with an integration over time (i.e., the mean of the ensemble is equal to the average

[1]This connection appears problematic and will be discussed in the general Conclusions.

over time), however

$$H(\psi) = -\int_{\psi} p(\psi) \ln p(\psi)\ \mathrm{d}\psi = -\frac{1}{T}\int_{0}^{T} p(\psi) \ln p(\psi)\ \mathrm{d}t \tag{3.2}$$

The free energy principle attempts to capture the existence of living systems in terms of global, pullback attractors in ergodic random dynamical systems with trajectories not undergoing phase transitions that would disrupt an agent, changing its description or worse, causing its death (Friston and Ao, 2012; Friston, 2012; Colombo and Wright, 2018). An intuitive example of this idea comes from homeostatic regulation of living systems and the maintenance of their essential variables within boundaries (Ashby, 1957). Following the ergodic assumption in equation (3.2), the (differential) entropy is now equal to the time average of surprisal, or self-information:

$$-\ln p(\psi) \tag{3.3}$$

All the measures defined so far omit, for simplicity, a conditional on the probability of states ψ: in the more general, Bayesian, case one should in fact write them as depending on a model or agent m (Robert, 2007; Friston, Thornton, and Clark, 2012; Barto, Mirolli, and Baldassarre, 2013). Intuitively, agents can share the same observations ψ but the description of the associated surprisal should be different: conditions of low surprisal for a fish, which express the set of states for its survival such as being in water, are not the same for a tiger, an oak tree or a bird, see also Kolchinsky and Wolpert (2018) for a review on ideas regarding the semantics or value of information. The more complete form of surprisal is thus defined as:

$$-\ln p(\psi|m) \tag{3.4}$$

but the conditional on m will mostly be omitted to simplify the notation. Under the free energy principle, biological systems must minimise their sensory surprisal over time. This minimisation is however intractable in any practical scenario, surprisal can in fact be seen as the negative log-model evidence or negative log-marginal likelihood of observations ψ, where the marginal likelihood is defined as:

$$p(\psi) = \int_{\vartheta} p(\psi, \vartheta)\ \mathrm{d}\vartheta \tag{3.5}$$

This integral is defined over all possible hidden (i.e., unobserved) states, inputs (sometimes referred to as "causes") and parameters/hyperparameters ϑ of observations ψ (see table 3.1). In many cases, the marginalisation is intractable since the latent space of ϑ can be high dimensional or the distribution can have a complex (analytical) form. In statistical mechanics, an approximation under variational formulations transforms this into an optimisation problem. The approximation goes by several names, including variational Bayes and ensemble learning (Hinton and Zemel, 1994; Dayan et al., 1995; Beal, 2003; MacKay, 2003; Wainwright, Jordan, et al., 2008; Bishop, 2006), and constitutes the mathematical basis of the free energy principle. Using variational Bayes, surprisal can be decomposed into (Bishop, 2006):

$$-\ln p(\psi) = F - D_{\mathrm{KL}}(q(\vartheta)\,||\,p(\vartheta|\psi)) \tag{3.6}$$

where

$$D_{\mathrm{KL}}(q(\vartheta)\,||\,p(\vartheta|\psi)) = \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta|\psi)}\,\mathrm{d}\vartheta \tag{3.7}$$

and

$$F \equiv \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta,\psi)}\,\mathrm{d}\vartheta \tag{3.8}$$

In equation (3.7) we define the Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951; Bishop, 2006), also known as relative entropy or information gain, an asymmetrical non-negative measure of the difference between two probability distributions. In this case, the two distributions are $p(\vartheta|\psi)$, the posterior distribution specifying the probability of hidden states, inputs and (hyper)parameters ϑ given observations ψ, and $q(\vartheta)$, a variational or recognition density. The latter, $q(\vartheta)$, is introduced with the idea of approximating the former, $p(\vartheta|\psi)$, which is defined using Bayes theorem and therefore depends on the marginal likelihood, making it intractable. This approximation entails the use of a simpler, known distribution $q(\vartheta)$ in order to minimise the difference between $q(\vartheta)$ and $p(\vartheta|\psi)$ through the KL divergence: when the difference is zero (the divergence is always non-negative), $q(\vartheta)$ is a perfect description of $p(\vartheta|\psi)$. From the point of view of an agent, the process is proposed to be analogous: to explain the hidden states, inputs and (hyper)parameters of sensations, $p(\vartheta|\psi)$, by approximating the posterior with a known distribution, $q(\vartheta)$. The other term, F, is defined in equation (3.8) and for its mathematical analogies with Helmholtz free energy in thermodynamics, it is named *variational* free energy (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston, 2008a; Friston et al., 2010b; Buckley et al., 2017), or (negative) evidence lower bound in machine learning (Bishop, 2006). Following Jensen's inequality, specifying that the KL divergence is always non-negative (Bishop, 2006), we get that:

$$D_{\mathrm{KL}}(q(\vartheta)\,||\,p(\vartheta|\psi)) \geq 0 \Rightarrow F \geq -\ln p(\psi) \tag{3.9}$$

thus defining variational free energy as an upper bound to surprisal, since by minimising F we are guaranteed to minimise $-\ln p(\psi)$. The divergence $D_{\mathrm{KL}}(q(\vartheta)||p(\vartheta|\psi))$ expresses the fact that the more we know about hidden states, inputs and (hyper)parameters ϑ of observations ψ, the more we minimise the surprise of observations ψ since good approximations $q(\vartheta)$ to the true posterior $p(\vartheta|\psi)$ make the divergence go to zero. If we knew exactly the hidden states, inputs and (hyper)parameters ϑ, such that $q(\vartheta) = p(\vartheta|\psi)$, then $D_{\mathrm{KL}}(q(\vartheta)\,||\,p(\vartheta|\psi)) = 0$ and following equation (3.8), we would have direct access to the surprisal since $-\ln p(\psi) = F$.

The free energy F can also be rewritten in two alternative forms, providing different interpretations on its minimisation. The first one specifies F as

$$\begin{aligned} F &= \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta, \psi)} \, \mathrm{d}\vartheta = \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta)} \, \mathrm{d}\vartheta - \int q(\vartheta) \ln p(\psi|\vartheta \, \mathrm{d}\vartheta) \\ &= D_{\mathrm{KL}}(q(\vartheta) \,||\, p(\vartheta)) - \mathbb{E}_{\mathrm{q}} \big[\ln p(\psi|\vartheta) \big] \end{aligned} \tag{3.10}$$

where the first term represents a measure of model complexity (cf. "Occam factor" (MacKay, 2003)) under a variational approximation generalising indexes such as the Bayesian information criterion (BIC). With Gaussian/Laplace assumptions, as we will see later, this complexity term reduces to the number of degrees of freedom of a model (Bishop, 2006). This definition is also related to Bayesian surprise (Itti and Baldi, 2009), as the difference between the variational density (i.e., the approximate posterior, cf. the real posterior in Bayesian surprise) and the prior on hidden states, inputs and parameters/hyperparameters, or the degrees of freedom of a system adopting a specific variational density. The second term expresses an accuracy measure in terms of the (negative) expected log-likelihood of the observations ψ. Under this decomposition, we can see that minimising variational free energy is equivalent to minimising the complexity of a model while maximising its accuracy (Daunizeau, 2017). The second form expresses variational free energy as

$$\begin{aligned} F &= \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta, \psi)} \, \mathrm{d}\vartheta = \int q(\vartheta) \big(- \ln p(\vartheta, \psi) \big) \, \mathrm{d}\vartheta + \int q(\vartheta) \ln q(\vartheta) \, \mathrm{d}\vartheta = \\ &= \mathbb{E}_{\mathrm{q}} \big[- \ln p(\psi, \vartheta) \big] - H(q(\psi)) \end{aligned} \tag{3.11}$$

where the last term is the negative Shannon entropy of the observations and the first one can be seen, in analogy with formulations in thermodynamics (see also the MaxEnt principle in Jaynes (1957a) and Jaynes (1957b)), as an "energy" term averaged over the variational density $q(\vartheta)$. This last formulation is then usually adopted for the definition of process theories that can provide algorithmic implementations for the minimisation of free energy, such as predictive coding (Friston and Kiebel, 2009c) and active inference (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston, 2008a; Buckley et al., 2017).

3.1.1 The generative density

To evaluate the variational free energy F for a system, we must formalise a generative density $p(\vartheta, \psi)$ and a recognition density $q(\vartheta)$ specific to an agent. Starting from the former, we define a generative model formulated as a one dimensional *generalised* state-space model (Friston, 2008a):

$$
\begin{aligned}
\dot{x} &= x' = f(x,v) + w \\
\dot{x}' &= x'' = f_x(x,v)x' + f_v(x,v)v' + w' \\
\dot{x}'' &= x''' = f_x(x,v)x'' + f_v(x,v)v'' + w'' \\
&\vdots \\
\psi &= g(x,v) + z \\
\psi' &= g_x(x,v)x' + g_v(x,v)v' + z' \\
\psi'' &= g_x(x,v)x'' + g_v(x,v)v'' + z'' \\
&\vdots
\end{aligned}
\tag{3.12}
$$

where ψ are the observations and where we expand $\vartheta = \{x, v, \theta, \gamma\}$, defining x as the hidden states and v as the exogenous inputs, while θ and γ, parameters and hyperparameters of $g(\cdot)$ and $f(\cdot)$, are still not explicitly represented here for simplicity and introduced later. Functions $g(\cdot)$ and $f(\cdot)$ map hidden states/inputs to observations and the dynamics of hidden states/inputs respectively. The prime signs, e.g., x', x'', x''' are used to define the generalised state of a variable. Generalised coordinates of motion are introduced to represent more general non-Markovian continuous stochastic processes (Stratonovich, 1967; Jazwinski, 1970; Friston, 2008a). Non-Markovian processes drop the assumption made in Ito's formulation of stochastic calculus, where the autocorrelation of a stochastic process can be seen as a delta function, representing discrete jumps (Jazwinski, 1970; Fox, 1987). Active inference schemes on the other hand, adopt a Stratonovich interpretation (Stratonovich, 1967), with smooth, continuous stochastic variables having finite, non-zero autocorrelation functions. In this light, random processes can be described as analytic (i.e., differentiable) and become better approximations of real-world (weakly) coloured noise (Fox, 1987; Van Kampen, 1992; Klöden and Platen, 1992; Friston, 2008a). With this formulation, it is then possible to define higher "orders of motion", or embedding orders, that encode, altogether, a trajectory for each of the variables used in this formulation. One could think of them as quantities conveying information about "velocity" (e.g., $(\cdot)'$), "acceleration" (e.g., $(\cdot)''$), etc. for each variable. For practical purposes, in equation (3.12) we've also taken a local linearity approximation on higher embedding orders suppressing nonlinear terms as suggested in (Friston, 2008a; Buckley et al., 2017) (i.e., ignoring high-order derivatives such as $g_{xx}(x,v), f_{xx}(x,v), g_{xxx}(x,v), f_{xxx}(x,v), \ldots$). We introduce then a more compact form:

$$
\begin{aligned}
\tilde{x}' &= f(\tilde{x}, \tilde{v}) + \tilde{w} \\
\tilde{\psi} &= g(\tilde{x}, \tilde{v}) + \tilde{z}
\end{aligned}
\tag{3.13}
$$

where the tilde sign (e.g., $\tilde{\psi}$) summarises the generalised state, a variable and its higher orders of motion (e.g., $\tilde{\psi} = \{\psi, \psi', \psi'', \dots\}$). The stochastic model in equation (3.12) can then be described in terms of a generative density where we make now explicit the dependence on variables θ, γ:

$$P(\tilde{\psi}, \tilde{x}, \tilde{v}, \theta, \gamma) = P(\tilde{\psi}|\tilde{x}, \tilde{v}, \theta, \gamma) P(\tilde{x}, \tilde{v}, \theta, \gamma). \tag{3.14}$$

$P(\psi|x, v; \theta, \gamma)$ is a likelihood function describing the measurement law in equation (3.13), while the prior $P(\tilde{x}, \tilde{v}; \theta, \gamma)$ describes the system's dynamics. In this formulation, we assume that parameters θ and hyperparameters γ can, in the most general case, be represented as time-varying quantities on a slower time scale, defining θ as parameters coupling hidden states and inputs to hidden states dynamics and observations, and hyperparameters γ as encoding properties of random fluctuations $\tilde{w}$ and noise $\tilde{z}$. In practice, it is often assumed that these quantities are (nearly) static variables, allowing for equation (3.14) to be expressed as

$$P(\tilde{\psi}, \tilde{x}, \tilde{v}; \theta, \gamma) = P(\tilde{\psi}|\tilde{x}, \tilde{v}; \theta, \gamma) P(\tilde{x}, \tilde{v}; \theta, \gamma) \tag{3.15}$$

using semicolons to highlight the fact that θ, γ are not random variables and don't change over time (or at least not as quickly as $\tilde{x}, \tilde{v}$) and that the conditional dependence does not apply to these variables. For most of the example applications presented in this thesis, I will use the simplified formulation in equation (3.15), but in Chapter 6 it will be crucial to consider (slowly) time-varying hyperparameters as represented in equation (3.14).

3.1.2 The variational density

At this stage, Friston proposes three different mathematical formulations incorporating different combinations of two approximations on the variational density $q(\vartheta)$:

- a mean-field partition of the unknowns ϑ based on their time scale, short for rapidly changing hidden states x and inputs v, longer for fixed or slowly changing parameters and hyperparameters θ, γ (Friston, 2008a; Buckley et al., 2017) and
- a Laplace assumption constraining the form of the variational density to a Gaussian $q(\vartheta)$ (Beal, 2003; Friston et al., 2007).

These three approaches go by the names of Dynamic Expectation Maximisation (DEM) (Friston, Trujillo-Barreto, and Daunizeau, 2008), Variational filtering (VF) (Friston, 2008b) and Generalised filtering (GF) (Friston et al., 2010b).

Dynamic Expectation Maximisation (DEM)

Dynamic Expectation Maximisation (Friston, Trujillo-Barreto, and Daunizeau, 2008) is a general variational framework for the estimation of unknown variables in a dynamical context. DEM combines a mean-field approximation on the set of unknown variables and a Laplace approximation of the generative density $p(\psi, \vartheta)$ using a Gaussian form for the variational density $q(\vartheta)$ (Beal, 2003; Friston et al., 2007). With a mean-field partition of variables based on a separation of different time-scales, the set of slowly changing parameters and hyperparameters are treated as conditionally independent with respect to hidden states and inputs of a stochastic dynamical system. In this light, while hidden states and inputs minimise the variational free energy, parameters and hyperparameters are assumed to minimise the path integral of the free energy since they are "fixed" for a long time scale. In DEM, this path integral is explicitly calculated and its results stored, with estimates of hidden states and inputs computed after each new observation, while parameters and hyperparameters are optimised only after a suitable amount of steps have been integrated and enough (model) evidence accumulated (Friston, Trujillo-Barreto, and Daunizeau, 2008).

The standard Laplace assumption approximates a distribution with a Gaussian density, meaning that only mean and variance (for a univariate case, otherwise the covariance matrix) need to be determined since they constitute a set of sufficient statistics. By finding the MAP estimate (i.e., the mode of a distribution), equal to the mean for Gaussian densities, and using a Taylor expansion around the mode, the variance is then analytically determined as the Hessian of an energy function (MacKay, 2003) evaluated at the mode (MacKay, 2003; Beal, 2003; Bishop, 2006; Friston et al., 2007; Särkkä, 2013). Traditionally, the Laplace method has been used in machine learning and statistics in an "offline" fashion, i.e., by first finding the mode of an available set of observations (ψ), and only then applying the approximation. The Laplace method generates a Gaussian distribution (rather than a point estimate, like MLE or MAP (Beal, 2003)) approximating an interval around the point with highest density of a distribution, i.e., its mode. Since it relies on MAP estimates however, it also suffers from the same weaknesses, mainly the problem that the mode may be representing the peak but not the mass of a distribution (Beal, 2003; Bishop, 2006). In DEM, the Laplace approximation is applied "online", fitting a Gaussian distribution around the local estimate of the mode updated over time while new observations are introduced (Penny, Kiebel, and Friston, 2006; Friston et al., 2007; Buckley et al., 2017). This allows the tracking of the mode of a distribution encoding time-varying dynamics, see the case for *active* inference where the distribution of observations ψ can be updated by an agent through action. With the optimal variance computed analytically, the minimisation of variational free energy is then reduced to a process of minimisation with respect to means/modes of the Laplace-approximated free energy updated, usually, via gradient descent (Friston, Trujillo-Barreto, and Daunizeau, 2008). However, this application of the Laplace assumption remains somewhat unclear in Friston's work, since the Laplace method is a good approximation only around the peak (MAP) of a distribution (MacKay, 2003; Bishop, 2006), a condition met only when the free energy is minimised by finding the optimal mean/mode of the approximate distribution. The "online" approximation proposed by Friston is, in this sense, not as clearly defined as the more traditional applications of the Laplace method, and its behaviour is less clear during its updates over time and away from equilibrium (see also section 3.1.3).

DEM can be seen as an extension of the Expectation-Maximisation (EM) algorithm (Baum and Eagon, 1967; Dempster, Laird, and Rubin, 1977) to handle stochastic rather than deterministic (i.e., known) dynamics with no fluctuations, and of Kalman filtering (Kalman, 1960b; Kalman and Bucy, 1961) to extend the estimation process to inputs (inputs), parameters and hyperparameters in generalised coordinates of motion with no Markov assumptions on random variables. DEM approaches are thought to generalise also the combined efforts of Kalman filtering and parameters estimation through EM such as the ones proposed by Roweis and Ghahramani (2001) and Beal (2003) with the introduction of generalised coordinates of motion for the treatment of coloured noise, although in principle one can re-write coloured noise as an auto-regressive (AR) process with an explicit treatment of stochastic variables still based on white noise where Kalman filtering theory applies (Friston, 2008a; Chui and Chen, 2017).

Variational filtering (VF)

Variational filtering generalises DEM by dropping the Laplace assumption, and thus extending the variational approximation to multimodal distributions in a dynamic context among others. VF gives a different perspective on stochastic approaches to the approximation of the posterior distribution $p(\vartheta|\psi)$. These approaches, usually based on sequential Monte Carlo sampling procedures, or particle filters, approximate a distribution by sampling a set of "particles" or candidate solutions from an approximate distribution of the posterior (Chen, 2003; Doucet and Johansen, 2009). These particles are then used to estimate the expected value of the observations (for each particle) and by comparing those expectations to the true observations. Parameters of the candidate approximate distribution are then updated based on the deviation or error of these particles from true observations. The most popular implementations also include a re-sampling step where particles with a low probability are regularly replaced by new ones from an updated version of the approximate posterior. By introducing a formulation in generalised coordinates of motion, VF drops the Markov assumption of particle filters, constraining the particles trajectories in the phase space with assumptions on higher embedding orders. This ensemble of candidate solutions then converges towards the true solution under the assumption that the dynamics of the observations ψ are slower than the ones of the optimisation process (i.e., a gradient descent procedure) (Friston, 2008b). The re-sampling phase, often used for practical purposes of fast convergence in particle filters, is thus unnecessary in variational filtering approaches since trajectories of the particles are better guided by constraints on their generalised motion (Friston, 2008b). As its Markovian counterpart however, VF is a computationally expensive procedure, which makes it potentially implausible for biological implementations, however see Sanborn and Chater (2016) for a different perspective on sampling methods in the brain.

Generalised filtering (GF)

Generalised filtering maintains a Laplace approximation of the posterior but drops the mean-field approximation based on time scales separation, allowing for conditional dependence between parameters/hyperparameters and hiddens states/inputs (Friston et al., 2010b). The main difference with DEM is that the integration of the variational free energy necessary to accommodate the minimisation of "fixed" parameters and hyperparameters is carried out implicitly in this case. Rather than minimising slowly changing variables by accumulating evidence over time and optimise these variables after a number of observations, GF approximates this process via a second order minimisation with respect to free energy, rather than a first order one on its path integral. The main assumption is that the variational free energy is smooth for changes of parameters and hyperparameters (Friston et al., 2010b), in some ways similar to the separation of times scales of DEM but less restrictive (e.g., no offline phase). For the derivation of an algorithmic implementation of variational free energy minimisation, i.e., active inference, in the remainder of the thesis I will focus on approaches implementing a Laplace approximation of the variational density, DEM and GF. The differences will be minimal until the optimisation of parameters and hyperparameters is considered (see Chapter 6), and in that case my agent-based models will make use of the GF formulation.

3.1.3 The Laplace assumption

Following the derivation in Buckley et al. (2017), we introduce a variational Gaussian approximation (Opper and Archambeau, 2009) and assume that the recognition density, $q(\vartheta)$, is Gaussian with mean μ and variance ς^2

$$q(\vartheta) \equiv \mathcal{N}(\vartheta; \mu, \varsigma^2) = \frac{1}{Z} \exp(\mathcal{E}) \tag{3.16}$$

where

$$Z \equiv \sqrt{2\pi\varsigma^2} \qquad \mathcal{E}(\vartheta) \equiv \frac{(\vartheta - \mu)^2}{2\varsigma^2} \tag{3.17}$$

After defining the *variational energy* (Friston et al., 2007)

$$E(\vartheta, \psi) \equiv -\ln p(\vartheta, \psi) \tag{3.18}$$

the variational free energy in equation (3.11) can be expressed as

$$F = \int q(\vartheta)\big(-\ln p(\vartheta,\psi)\big)\ \mathrm{d}\vartheta + \int q(\vartheta)\ln q(\vartheta)\ \mathrm{d}\vartheta = \int q(\vartheta)\big(E(\vartheta,\psi)\big)\ \mathrm{d}\vartheta + \int q(\vartheta)(-\ln Z - \mathcal{E}(\vartheta))\ \mathrm{d}\vartheta \tag{3.19}$$

where, for simplicity, we momentarily dropped the use of generalised coordinates, i.e., $\tilde{\vartheta}, \tilde{\psi}$. The second integral in equation (3.19) can be evaluated analytically, in particular the first part reduces to

$$\int q(\vartheta)(-\ln Z)\ \mathrm{d}\vartheta = -\int q(\vartheta)\ln\left(\sqrt{2\pi\varsigma^2}\right)\ \mathrm{d}\vartheta = -\ln\left(\sqrt{2\pi\varsigma^2}\right) = -\frac{1}{2}\ln(2\pi\varsigma^2) \tag{3.20}$$

since the factor Z does not depend on ϑ and the integral of the variational density $q(\vartheta)$ over ϑ is 1. The second term of the second integral can be manipulated into

$$\int q(\vartheta)(-\mathcal{E}(\vartheta))\ \mathrm{d}\vartheta = \int q(\vartheta)(-\frac{(\vartheta-\mu)^2}{2\varsigma^2})\ \mathrm{d}\vartheta = -\frac{1}{2\varsigma^2}\int q(\vartheta)(\vartheta-\mu)^2\ \mathrm{d}\vartheta = -\frac{1}{2} \tag{3.21}$$

and the integral is in this case equivalent to the definition of the variance ς.

According to the traditional Laplace or saddle-point (MacKay, 2003) method, to solve the first integral in equation (3.19), one approximates the generative density with a Taylor expansion around its MAP estimate $\hat{\vartheta}$ up to second order (i.e., a Gaussian approximation) (Beal, 2003)

$$E(\vartheta,\psi) \approx E(\hat{\vartheta},\psi) + (\vartheta-\hat{\vartheta})\ \frac{\partial E}{\partial \vartheta}\bigg|_{\vartheta=\hat{\vartheta}} + \frac{(\vartheta-\hat{\vartheta})^2}{2}\ \frac{\partial^2 E}{\partial \vartheta^2}\bigg|_{\vartheta=\hat{\vartheta}} \tag{3.22}$$

where terms of order $O(\vartheta^3)$ are dropped. If we assume that the MAP estimate is known, then the variational density $q(\vartheta)$ will be centred around this peak (Bishop, 2006),

$$\hat{\vartheta} = \mu \tag{3.23}$$

I will use μ from now on for consistency with work such as Friston et al. (2007), Daunizeau (2017), and Buckley et al. (2017), in place of $\hat{\vartheta}$, used in the more traditional statistics/machine learning literature to represent MLE/MAP estimates. We can then replace this expression into the first integral of equation (3.19), giving

$$\begin{aligned}
&\int q(\vartheta)\big(E(\vartheta,\psi)\big)\,\mathrm{d}\vartheta = \\
&= \int q(\vartheta)\left(E(\mu,\psi) + (\vartheta-\mu)\left.\frac{\partial E}{\partial \vartheta}\right|_{\vartheta=\mu} + \frac{1}{2}(\vartheta-\mu)^2\left.\frac{\partial^2 E}{\partial \vartheta^2}\right|_{\vartheta=\mu}\right)\mathrm{d}\vartheta = \\
&= E(\mu,\psi) + \left.\frac{\partial E}{\partial \vartheta}\right|_{\vartheta=\mu}\int q(\vartheta)(\vartheta-\mu) + \frac{1}{2}\left.\frac{\partial^2 E}{\partial \vartheta^2}\right|_{\vartheta=\mu}\int q(\vartheta)(\vartheta-\mu)^2
\end{aligned} \tag{3.24}$$

and since we are considering a MAP estimate $\hat{\vartheta}$ ($=\mu$) of $E(\vartheta,\psi)$, the first derivative of the variational energy evaluated at the MAP estimate is zero

$$\left.\frac{\partial E}{\partial \vartheta}\right|_{\vartheta=\mu} = 0 \tag{3.25}$$

leaving only the first and the third term plus the expressions derived from solving analytically the negative entropy part of equation (3.19)

$$F = E(\mu,\psi) + \frac{1}{2}\left(\left.\frac{\partial^2 E}{\partial \vartheta^2}\right|_{\vartheta=\mu}\varsigma^2 - \ln(2\pi\varsigma^2) - 1\right) \tag{3.26}$$

where we used once again the definition of the variance ς^2 to simplify the last term of the Taylor expansion. In the standard Laplace approximation (Azevedo-Filho and Shachter, 1994), one can then find an explicit form for the variance ς^2 in equation (3.26) by differentiating F with respect to ς^2 and obtain

$$\frac{\delta F}{\delta \varsigma^2} = 0 \implies \varsigma^2 = \left(\left.\frac{\partial^2 E}{\partial \vartheta^2}\right|_{\vartheta=\mu}\right)^{-1} = \bar{\varsigma}^2 \tag{3.27}$$

where $\bar{\varsigma}^2$ represents the *optimal* variance evaluated at the MAP estimated and only with respect to hyperparameters, $\vartheta = \{\gamma\}$. After replacing $\varsigma^2 = \bar{\varsigma}^2$, the variational free energy is then reduced to the sum of the variational energy evaluated at the mode of the generative density (i.e., the MAP estimate) and a term that can be calculated analytically using the variational energy $E(\mu,\psi)$ at the mode (and will thus be normally dropped in our implementations, making $F \approx E$)

$$F = E(\mu,\psi) - \frac{1}{2}\ln(2\pi\bar{\varsigma}^2) \approx E(\mu,\psi) = -\ln p(\mu,\psi) \tag{3.28}$$

The method implemented by Friston (Penny, Kiebel, and Friston, 2006; Friston et al., 2007) diverges in some ways from the traditional Laplace approximation. The proposal of an online version of the Laplace method (also called variational Laplace (Daunizeau, 2017)) implies that the MAP estimate is not readily available as assumed after equation (3.22) but rather, inferred as new data becomes available (e.g., an agent perceiving the world over time). This implies that equation (3.24) is true only at the MAP estimate (Daunizeau, 2017) and not necessarily accurate away from it, e.g., during the inference phase of the MAP estimate or if the peak of the generative density changes over time due to an agent's actions on the world. Friston et al. (2007) suggested that the over-reliance on MAP estimates (the main weakness of the Laplace approximation (Beal, 2003)) is bypassed by the online version thanks to a *factorised* variational formulation that fits a Gaussian to the generative density after each new data point is introduced and for each set of variables within ϑ in the case of a mean-field approximation. Rather than relying on a post hoc point estimate of the MAP of a density where all data is available, the online Laplace method is claimed to better approximate the mass of a distribution using MAP estimations of each set of variables as constraints for the other (MAP) estimates, instead of relying on the inference of a single mode for the generative density, see the mean-field-induced factorisation in Friston et al. (2007) and Daunizeau (2017). In generalised filtering however, these constraints are removed since the mean field approximation is dropped (Friston et al., 2010b), suggesting that the "variational" approximation (almost) reduces to a standard Laplace approximation once the variational free energy is minimised (Opper and Archambeau, 2009). In both cases, the optimal variance $\vec{\varsigma}^2$ described in equation (3.27) is accurate only at the peak of the generative density, requiring a more complex formulation away from it. As explicitly shown by Buckley et al. (2017), the derivation proposed by Friston et al. (2007) and Friston, Trujillo-Barreto, and Daunizeau (2008) simply drops the first derivative of the variational energy obtained from a Taylor expansion around the peak as shown in equation (3.25) (only true at the MAP estimate), following the traditional offline Laplace method. This may be due to the fact that the effects of the first term of the Taylor expansion are relatively small away from the peak, but it remains unclear how disruptive this assumption for the online Laplace approximation really is. Future work should thus address the details of this approach, crucial also for the SPM methodologies in neuroimaging analysis (Penny, Kiebel, and Friston, 2006). In this thesis I will simply assume that online Laplace is a good approximation of the offline version. To highlight the fact that this is a different variation of the Laplace method, I will however use a slightly different notation for equation (3.28):

$$F \approx -\ln p(\mu, \psi) = -\ln p(\vartheta, \psi)\Big|_{\vartheta=\mu} \tag{3.29}$$

showing that the approximate equality is exactly true only for the MAP estimate $\vartheta = \hat{\vartheta}(= \mu)$.

3.1.4 The Laplace-encoded variational free energy

Starting from the variational free energy in equation (3.28), expanding ϑ in terms of hidden states, inputs, parameters and hyperparameters, reintroducing generalised coordinate of motion and not including terms treated as constants during the minimisation process (i.e., the ones depending on the optimal variance $\bar{\varsigma}^2$) we get

$$F \approx -\ln p(\tilde{\psi}, \tilde{x}, \tilde{v}, \theta, \gamma)\Big|_{\tilde{\vartheta}=\tilde{\mu}} \tag{3.30}$$

where $\tilde{\vartheta} = \tilde{\mu}$ represents the fact that the generative density $P(\tilde{\psi}, \tilde{x}, \tilde{v}, \theta, \gamma)$ will be approximated by a Gaussian distribution centred around the MAP estimates $\hat{\tilde{\vartheta}} = \tilde{\mu}$ of $\tilde{\vartheta}$ following the online Laplace method explained above. Diverging from the extensive review found in Buckley et al. (2017), I will write an expression for the variational free energy based on a treatment of – all – unknown quantities ϑ as random variables approximated under the Laplace assumption. In Buckley et al. (2017), the authors use this assumption only on hidden states x and inputs v, essentially following a formulation based on DEM (Friston, Trujillo-Barreto, and Daunizeau, 2008), see also the paragraph on DEM above. This derivation simplifies the representation of variables θ and γ, reported effectively as fixed parameters in Buckley et al. (2017) and in much of the work presented here (Chapter 4, Chapter 5 and Chapter 7), but needs to be generalised for time-varying parameters and hyperparameters as in the case of the implementation presented in Chapter 6.

For convenience, from now on I will assume that random variables $\tilde{z}$ and $\tilde{w}$ are zero-mean Gaussian with precision matrices (inverse covariances) $\Pi_{\tilde{z}}, \Pi_{\tilde{w}}$

$$\begin{aligned} \tilde{z} &\sim \mathcal{N}(\tilde{z}; 0, \Pi_{\tilde{z}}) \\ \tilde{w} &\sim \mathcal{N}(\tilde{w}; 0, \Pi_{\tilde{w}}) \end{aligned} \tag{3.31}$$

in equation (3.13), meaning that the likelihood and prior in equation (3.14) are also Gaussian. Notice that in this case we defined precision *matrices* even for a univariate case because of the presence of extra embedding orders of motion that effectively increase the number of variables. Dropping momentarily the use of generalised coordinates for simplicity, we will assume that π_z and π_w represent elements on the main diagonals of their respective precision matrices. Likelihood and prior in equation (3.14) can then be explicitly written as

$$\begin{aligned} P(\psi|x, v, \theta, \gamma) &= \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(\frac{-\big(\psi - g(x, v, \theta)\big)^2}{2\sigma_z^2}\right) \\ P(x, v, \theta, \gamma) &= \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left(\frac{-\big(x' - f(x, v, \theta)\big)^2}{2\sigma_w^2}\right). \end{aligned} \tag{3.32}$$

When substituted in equation (3.30) and evaluated at $\vartheta = \mu$, i.e., $x = \mu_x, v = \mu_v, \theta = \mu_\theta, \gamma = \mu_\gamma$, the variational free energy can be expressed as:

$$F \approx \frac{1}{2}\left[\mu_{\pi_z}\big(\psi - g(\mu_x, \mu_v, \mu_\theta)\big)^2 + \mu_{\pi_w}\big(x' - f(\mu_x, \mu_v, \mu_\theta)\big)^2 - \ln(\mu_{\pi_z}\mu_{\pi_w})\right] \tag{3.33}$$

Expected precisions μ_π (or more in general $\mu_{\tilde{\Pi}}$) are used instead of μ_γ or $\tilde{\mu}_\gamma$ since the latter represent a simple re-parametrisation of precisions introduced later for the optimisation of these variables. Explicit precision-weighted prediction errors for θ, γ are not introduced, unlike Friston (2008a) where predictions specify also prior expectations on parameters and hyperparameters, to keep the formulation more manageable. For implementations in most of the later chapters, I will use θ, γ, highlighting that these parameters and hyperparameters are fixed quantities in my models, and only in Chapter 6 I will specifically shift to a representation of the sufficient statistics for hyperparameters $\mu_{\tilde{\gamma}}$ when they will be optimised.

It is also important to highlight that the assumption in equation (3.31), even if common in derivations of the FEP (see for instance Friston, Trujillo-Barreto, and Daunizeau (2008) and Buckley et al. (2017)), de facto renders the variational approximation exact. Since both the variational and generative densities are now Gaussian, the variational approach simply transforms a problem of inference into an optimisation problem (Daunizeau, 2017). One could thus, in principle, find analytically the optimal variational density to approximate the generative density whenever the problem is tractable or use alternative, arguably simpler, methods. It may then look redundant to consider variational approaches. However, one of the strengths of the variational approach lies in the fact that one could define any form for the generative density and still use variational formulations/Laplace approximation for their computational efficiency (Bishop, 2006), even if the quality of the approximation may deteriorate quickly for increasingly complex (e.g., multimodal) distributions. The variational formulation would be the same for any distribution and could be implemented with hierarchical models using even discrete distributions at higher layers, see Parr and Friston (2018c) for example, acting as probability integral transforms for lower ones, thus extending its application beyond Gaussian distributions (Friston, 2008a). For these reasons, I will focus on variational formulations for their generality, even if simpler methods could in some cases be applied.

3.2 The minimisation of variational free energy

With the expression in equation (3.33), it is now possible to define explicitly the minimisation scheme for the variational free energy under the Laplace assumption. Given the set of unknowns ϑ, different processes can be implemented based on the minimisation of the free energy functional with respect to hidden states, inputs, parameters and hyperparameters, and most importantly action (introduced later). These processes are usually based on (stochastic) gradient descent procedures, finding the minimum of a function(al), the variational free energy in our case, by following the steepest direction defined by the gradient (Friston, Trujillo-Barreto, and Daunizeau, 2008). At the moment, the choice of using a gradient descent appears somewhat arbitrary and motivated mainly by the simplicity of the method (Kim, 2018). More in general one could apply other methods, either augmenting the gradient descent with extra terms such as in Adagrad, Adadelta, Adam, etc., see for instance Ahmadi and Tani (2018), or by using different approaches such as in Baltieri (2014). On the other hand, the Fokker-Planck formulation of solutions following simple Langevin dynamics proposed in Friston, Trujillo-Barreto, and Daunizeau (2008) and Friston (2008b) strongly resonates with work such as Mandt, Hoffman, and Blei (2016) and Chaudhari and Soatto (2018) in machine learning and Seifert (2012) in (stochastic) thermodynamics, suggesting that stochastic gradient descent methods can be seen as a process of approximate Bayesian inference themselves (and therefore optimal given different sets of constraints, or priors) if analysed from the perspective of variational calculus. In support of gradient descent methods, Jordan, Kinderlehrer, and Otto (1998) also proposed variational formulations for solutions to the Fokker Plank equation in statistical mechanics, essentially based on a steepest or gradient descent principle. In this thesis I will follow the gradient descent idea proposed by standard implementations of the FEP and discuss only this minimisation procedure from here onwards. As we will see, this choice is especially interesting since it maps to pre-existing methods including system identification techniques, estimation procedures, deconvolution (see Friston (2008a) for a review) and control (see Chapter 6).

This minimisation, implemented for different sets of variables, has been suggested to describe and account of cognitive processes, such as perception, action, learning and attention. Inspired by existing proposals regarding, for instance, perception as inference and action as optimal control, the minimisation of variational free energy as been proposed as a general computational principle of cognitive functions, as suggested from a review of the literature in Chapter 2.

3.2.1 Perception

Following Friston, Trujillo-Barreto, and Daunizeau (2008) and Friston (2008a), the optimisation of the (Laplace-encoded) variational free energy with respect to expected hidden states $\tilde{\mu}_x$ and expected inputs $\tilde{\mu}_v$ can be interpreted as estimation or perception:

$$\dot{\tilde{\mu}}_x = D\tilde{\mu}_x - \frac{\partial F}{\partial \tilde{\mu}_x} \tag{3.34}$$

$$\dot{\tilde{\mu}}_v = D\tilde{\mu}_v - \frac{\partial F}{\partial \tilde{\mu}_v} \tag{3.35}$$

These sets of equations (the tilde collapses all orders of motion to a single term for convenience, so each equation is in itself a system of equations) include extra terms $D\tilde{\mu}_x, D\tilde{\mu}_v$ to represent the "mode of the motion" (also the mean for Gaussian variables) in the minimisation of variables in generalised coordinates of motion (Friston, 2008a; Buckley et al., 2017; Kim, 2018), with D as a differential operator "shifting" the order of motion of $\tilde{\mu}_x, \tilde{\mu}_v$ such that $D\tilde{\mu}_x = \tilde{\mu}'_x$ and $D\tilde{\mu}_v = \tilde{\mu}'_v$. More intuitively, since we are now minimising the components of a generalised state representing a trajectory rather than a static state, variables are in a moving framework of reference in the state-space, and the minimisation is achieved for $\dot{\tilde{\mu}}_x = \tilde{\mu}'_x, \dot{\tilde{\mu}}_v = \tilde{\mu}'_v$ rather than for $\dot{\tilde{\mu}}_x = 0, \dot{\tilde{\mu}}_v = 0$ (which would assume that the mode of the motion is zero, as in standard state-space formulations with Markov assumptions). In all the implementations included in this thesis, expected inputs $\tilde{\mu}_v$ are not updated and considered as fixed for the purposes of building effective controllers. The presence of a minimisation with respect to $\tilde{\mu}_v$ highlights however the already mentioned generalisation of Kalman filters, which in their most basic form are unable to recover inputs or " causes" from observations. For specific discussions on this last point see Friston, Trujillo-Barreto, and Daunizeau (2008) where a comparison of the DEM algorithm and other approaches, including Kalman filtering, sheds light on their similarities and differences.

3.2.2 Action

In active inference, action is described as a problem of optimal control that essentially mirrors perception by changing observations $\tilde{\psi}$ to better match expected hidden states $\tilde{\mu}_x$. This process is based on the general assumption that, from the perspective of an agent, observations ψ are affected by actions a (ψ is a function of a, $\psi(a)$), and that this is the only thing an agent can be sure of. Action is thus cast as:

$$\dot{a} = -\frac{\partial F}{\partial a} = -\frac{\partial F}{\partial \tilde{\psi}} \frac{\partial \tilde{\psi}}{\partial a} \tag{3.36}$$

This assumption is proposed in order to bypass a well known problem of inverse models for motor systems and control theory more generally, the redundancy of effective movements (Bernstein, 1967; Sporns and Edelman, 1993; Franklin and Wolpert, 2011). Finding the action or policy (a sequence of actions) responsible for some observation by inverting a forward model (Kawato, 1999; Wolpert and Ghahramani, 2000), i.e., working out the effects of the action/policy on hidden states and inputs generating the observations, can easily generate an ill-posed problem: a one-to-many mapping between a single observation and the set of all possible actions that could have generated. In active inference, this inverse model is decomposed into two sub-problems based on the presence of an intrinsic (bodily) and an extrinsic (environmental) frames of reference (Friston, 2011). In the intrinsic frame, it is proposed that a large portion of the control problem can be solved by predicting proprioceptive sensations[2] in the same way exteroceptive[3] ones are, by using observations and a generative model of their dynamics to generate estimates of their state. The

[2]Proprioception is the sense of position and movement of different parts of our body, as seen in Chapter 4.

[3]Exteroception captures perceptual modalities at the interface with the environment: vision, touch, olfaction, taste and audition.

problem in extrinsic coordinates is solved by finding simple, heuristic mappings between proprioceptive estimates and observations that can be trivially implemented as low-level reflexes. With this separation, the heavy lifting is performed by a generative model, now also producing predictions of proprioceptive states, while the extrinsic problem in the form of reflex arcs is thought to be "solved" by agents over an evolutionary scale (Friston et al., 2010a), in the same way reflexes are not "learnt" during development but are assumed to be developed by a phylogeny over several generations. This interpretation is heavily reliant on proprioception, often implicitly studied and assumed to exist only in complex life forms. However, the idea of simple reflexes driving behaviour can also be relevant to simpler organisms, with "proprioceptive predictions" that could be ascribed to even simple chemical networks triggering a reflex, such as tumbling in bacteria. For example in Yi et al. (2000) and Andrews, Yi, and Iglesias (2006), the chemical network is effectively implementing an integral control mechanism, which I derive in terms of active inference in Chapter 6.

One could also argue that equation (3.36) is still fundamentally an inverse model (i.e., finding the correct action for a desired output), but unlike more traditional approaches, active inference does not involve a mapping from hidden states $\tilde{x}$ to actions a, since it is cast in terms of sensory data $\tilde{\psi}$ directly. The problem is thus reduced from a mapping between *unknown* hidden states and actions, to a mapping between *known* observations $\tilde{\psi}$ and actions a, see Fig. 16 in Friston et al. (2010a). This implementation also resonates with sensorimotor accounts of agent-environment systems where action is fundamentally grounded in an extrinsic frame of reference (Buhrmann and Di Paolo, 2014), i.e., the real world ($\tilde{\psi}$), rather than in a intrinsic one in terms of hidden states ($\tilde{x}$) to be inferred first by inverting an internal forward model. The connection to classic frameworks of optimal control can also be readily recovered if one assumes that knowledge of the mapping between actions a and hidden states x is known, rewriting equation (3.36) as:

$$\dot{a} = -\frac{\partial F}{\partial a} = -\frac{\partial F}{\partial \tilde{\mu}_x}\frac{\partial \tilde{\mu}_x}{\partial a} \tag{3.37}$$

3.2.3 Learning

Learning is described as the minimisation of the path integral of variational free energy, or free action, with respect to parameters θ, encoding slowly changing properties of the coupling between hidden states/inputs and hidden states dynamics/observations (Friston, 2008a; Friston, 2010b; Buckley et al., 2017). For processes of learning, one can assume that parameters θ are fixed using a mean field approximation, see DEM (Friston, Trujillo-Barreto, and Daunizeau, 2008) and paragraph above. Alternatively, these parameters can be represented as slowly changing (random variables) with respect do hidden states and inputs even if a formal separation of time scales is not assumed (no mean field approximation, see GF (Friston et al., 2010b) and above).

In this case I will proceed by applying the methodology proposed by Generalised Filtering, specifying instantaneous changes on the curvature of expected parameters with respect to variational free energy (i.e., GF) rather than first order updates with respect to explicit calculations of free action (i.e., DEM):

$$\ddot{\mu}_\theta = -\frac{\partial F}{\partial \mu_\theta} \tag{3.38}$$

This equation is presented for completeness but in the remainder of the thesis never effectively used since all of my work mainly focuses on action and perception processes and only in one case on the role of precision hyperparameters optimisation presented next.

3.2.4 Attention

Attention mechanisms (Friston, 2010b; Feldman and Friston, 2010) are suggested to be closely related to the optimisation of precisions π of both sensory and system random variables z and w. In equation (3.33), these precisions effectively modulate different prediction errors. "Attention" is in this case thought to be the process that regulates the optimisation of other variables (hidden states, inputs and parameters) by updating weights in the Laplace-encoded free energy reduced to a weighted sum of different prediction error terms. "Attending" to some information, perhaps one modality (e.g., vision) over a second one (e.g., proprioception) is, according to this hypothesis, represented by an increase of precisions of the former, a decrease for the second, or a combination of both cases, see Feldman and Friston (2010) and Chapter 4 for discussion. This process is implemented with the same assumption on the time scale highlighted for parameters, where a second order scheme is introduced following GF (Friston et al., 2010b):

$$\ddot{\mu}_\pi = -\frac{\partial F}{\partial \mu_\pi} \tag{3.39}$$

Hyperparameters (i.e., expected precisions) μ_π should however be non-negative since variances need to be positive. To include this constraint, following Friston et al. (2007) we thus parametrise precisions π (and consequently expected precisions μ_π) in the generative model as:

$$\pi = \exp(\gamma) \tag{3.40}$$

creating log-normal priors on (expected) sensory precisions and making them strictly positive thanks to the exponential mapping of hyperparameters γ (as part of the previously defined ϑ). The scheme in equation (3.39) is thus replaced by one in terms of expected log-precisions μ_γ:

$$\ddot{\mu}_\gamma = -\frac{\partial F}{\partial \mu_\gamma} \tag{3.41}$$

TABLE 3.1: **The variables used in the derivation of the FEP.**

Expression	Description
m	Model or agent
ψ	Sensory input/observations/measurements
ϑ	Set of unobserved, or hidden, variables used in the generative model, $\vartheta = \{x, v, \theta, \gamma\}$
x	Hidden states of the generative model
v	Inputs/causes of the generative model
θ	Parameters of the generative model
γ	Hyperparameters of the generative model, used to map precisions/variances to log-normal (strictly positive) values
a	Actions of an agent with effects on world dynamics described by the generative process
$\tilde{\psi}, \tilde{x}, \tilde{v}$	Generalised variables, a variable and its higher orders of motion (i.e., embedding orders), e.g., $\tilde{\psi} = \{\psi, \psi', \psi'', \dots\}$
$-\ln(\psi \| m)$	Surprisal or self-information, measuring how unlikely an observation ψ is for an agent m
$H(\psi)$	Shannon entropy or average/expected surprisal of observations ψ
$D_{KL}(p\|\|q)$	Kullback-Leibler (KL) divergence, an asymmetrical measure of the difference between two distributions p and q
F	Variational free energy functional
$p(\psi, \vartheta)$	Generative density, an agent's description of its observations
$q(\vartheta)$	Recognition or variational density, used to approximate the generative density in Variational Bayes
$g(x, v)$	Function mapping from inputs and hidden states to observations
$f(x, v)$	Function mapping from inputs and hidden states to hidden states dynamics

Expression	Description
$\tilde{z}$	Random variables used to represent analytical (i.e., differentiable) noise on observation/measurement uncertainty
$\tilde{w}$	Random variables used to represent analytical (i.e., differentiable) random fluctuations/uncertainty in system dynamics
$\Pi_{\tilde{z}}$	Precision (inverse covariance) matrix of $\tilde{z}$
$\Pi_{\tilde{w}}$	Precision (inverse covariance) matrix of $\tilde{w}$
$E(\vartheta, \psi)$	Variational energy, used for the Laplace approximation
Z	Normalising factor or partition function, used for the Laplace approximation
$\mathcal{E}(\vartheta)$	Energy of variables ϑ, in analogy to Boltzamnn distribution
μ	Means/modes of the variational density for the Laplace approximation
ς	Variance of the variational density for the Laplace approximation
$\hat{\vartheta}$	MAP estimate of the generative density
$\mu_x, \mu_v, \mu_\theta, \mu_\gamma$	Expectations on variables x, v, θ, γ after the use of the Laplace assumption
D	Differential operator shifting the embedding order by one for consistency in minimisation of variational free energy in generalised coordinates of motion

Chapter 4

A simple action-perception loop in active inference

In active inference, perception is a process of inferring the causes of sensory data by minimising the error between actual sensations and those predicted by a probabilistic generative model. Action, on the other hand, is specified as a process that modifies the world such that the consequent sensory input meets expectations encoded by the same internal model. These two processes, inferring properties of the world and inferring actions needed to meet expectations, close the sensorimotor loop and suggest a deep symmetry between action and perception. This chapter provides a minimal model of a sensorimotor loop under active inference formulations to illustrate its main features. I will introduce a simple cruise controller, an agent attempting to regulate its velocity on a flat surface and focus on showing different properties and working regimes of this agent, arising from the interaction of perception and action processes according to the FEP formulation. One of the main goals is to discuss the implementation of passive and active agents depicted in Chapter 1 (see also "perception-centric" and "action-oriented" models of cognition under the FEP in Chapter 2). By simply tweaking different parameters, i.e., precision weights, I'll show how the agent's behaviour changes in non-trivial ways. Most of the results reported here will be expanded in later chapters and therefore some of the implications will be discussed and elaborated there. This example should get the reader more familiar with a set of ideas used in the remainder of the models. One of the main assumptions of this thesis, i.e., the fact that an agent is not a copy-model or mirror of its environment, will be momentarily dropped in favour of clarity for our first implementation and also because the equations describing the world dynamics, as we will see, could hardly ever be simplified by an agent.

4.1 Action in an active inference context

The role of action in the minimisation of variational free energy under the FEP was recognised with the introduction of active inference in Friston, Daunizeau, and Kiebel (2009) and Friston et al. (2010a), suggesting that biological systems not only need to update predictions of their generative models through perception, they must also actively engage with the environment to minimise prediction errors by creating sensory input predicted by such models. This perspective, as highlighted in Chapter 2, is particularly relevant for this thesis since it provides some of the necessary tools to connect the FEP with 4E ideas highlighting the importance of behaviour, in particular sensorimotor loops and online feedback mechanisms.

However, the active inference proposal introduces several questions regarding the behaviour of systems minimising free energy when actively interacting with the world, e.g., Friston, Thornton, and Clark (2012). More specifically, the unifying perspective of perception and action as processes of inference using the same generative model (Friston et al., 2010a; Clark, 2013) presents some fundamental differences with some of the more traditional approaches to the modelling of sensorimotor loops. As discussed in more detail in Chapter 7, proposals accounting for these processes in terms of Bayesian inference often focus on either perception or action, expecting them to be sequentially (and optimally) combined following the traditional sense-model-plan-act approach (Brooks, 1991b) of traditional cognitive science/AI/robotics, e.g., Knill and Richards (1996), Todorov and Jordan (2002), Todorov (2004), and Franklin and Wolpert (2011). Inspired by similar ideas of action/control and planning as inference, e.g., Attias (2003), Botvinick and Toussaint (2012), and Kappen, Gómez, and Opper (2012), active inference also relies on the intuition that describes actions in terms of backward estimation starting from desired goals treated as "observed" states in the future (Todorov, 2009b). Unlike other approaches in control theory, however (cf. (Kawato, 1999; Wolpert and Ghahramani, 2000; Todorov, 2004), the estimation of hidden states and actions (n.b. also treated as latent variables) is "shared" within a single (generative) model: the sequential perception of external stimuli followed by the description of models of the environment later implementing motor commands is replaced by a single set of processes that perform all of these tasks in a more unified and cohesive way. Following the proposal advanced by active inference, an agent is engaging in behaviour driven by a mix of estimates of properties of the world and an agent's desires/priors that, following Bayes theorem, are combined into predictions triggering motor actions and enacting (normative) behaviour without the need for precise models of the world. Passive agents (see Chapter 1) that merely implement accurate estimates of hidden properties of the environment, can still be described by active inference (Allen and Friston, 2018) but provide no real explanation of adaptive behaviour (Bruineberg, Kiverstein, and Rietveld, 2018).

Crucial in the active inference scheme is the role played by *proprioception* (the sense of position and movement of different parts of our body), also constituting the focus of all the experiments presented in this thesis. Active inference suggests that an agent ought to produce its own (proprioceptive) predictions errors that can then be solved by action and implemented via reflex arcs (Friston, 2011; Brown et al., 2013). On this view, an agent should generate predictions of incoming sensations explicitly misaligned with world variables ("Action is enabled by systematic misrepresentations", Wiese (2016)). For instance, in order to grab a mug on a table a few steps away from me, active inference suggests that I should (strongly) predict my hand to be around the mug so that visual, tactile and especially proprioceptive sensations (my hand still being in my pocket) can generate prediction errors that are then solved by motor actions: taking a few steps towards the table, getting my hand out of my pocket and reaching for the mug. This turns the more traditional understanding of perception around: rather than estimating the hidden properties of the world responsible for our sensory input to then accurately model the environment and plan our next actions, active inference agents must explicitly mispredict their sensations in order to bring their desires into existence via simple error minimisation procedures resolved by appropriate motor actions. In this view, perception becomes de facto a description of both an agent's desires and world variables, mashed in a way that bears little value to anyone but the agent itself.

To show the role of both priors and precisions in the context of sensorimotor loops, in this chapter I present some initial results from computational simulations of active inference agents performing basic homeostatic control. By focusing on a minimal model of a "Bayesian cruise controller" similar in spirit to the example of a "Bayesian thermostat" found in Buckley et al. (2017), I will emphasise the role of perception and action in different settings of top-down and bottom-up processing achieved by simply modifying the relative strength of different precisions weights. These initial simulations will also reflect the emphasis of this thesis to work on minimal models. I will not consider, at the moment, more advanced constructs provided in other treatments of the FEP, including for instance hierarchies and generalised coordinates of motion (Friston, 2008a) (generalised coordinates will however be used in Chapter 6).

4.2 A Bayesian cruise controller

In this model, a block of mass = 1 kg (our agent) is placed on a surface with some sliding friction. The goal of this agent is to regulate its velocity, which can be perceived through a sensor, towards a desired set-point v_{des} ($v_{des} = 10$ m/s unless otherwise stated). The regulation will be described as a Bayesian inference process, inspired by the free energy principle and implemented in an active inference set up. The details behind the mechanism for velocity regulation will not be specified, since they don't add any more insight to our proof of concept. We will simply assume that this agent can apply a force that moves the block against the effects of friction which tend to bring the velocity of the block down to zero. The *generative process*, describing the dynamics of the world for our agent, will simply entail the definition of a velocity variable x (here to be interpreted as hidden state rather than as a position/displacement) that exponentially decays over time with a constant rate α due to the effects of friction. We also describe these dynamics as noisy, with a random variable $w \sim \mathcal{N}(0, \sigma_w^2)$, and have an action variable a that represents the force applied by the agent as an input (in states-space formulations terms) to achieve homeostatic control. The generative process is presented in the form of a state-space model as in most implementations of active inference, e.g., Friston, 2008a; Buckley et al., 2017; Bogacz, 2017; Baltieri and Buckley, 2019c:

$$x' = -\alpha x + a + w \tag{4.1}$$

To simplify the example, no other exogenous inputs (again in a state-space represen-

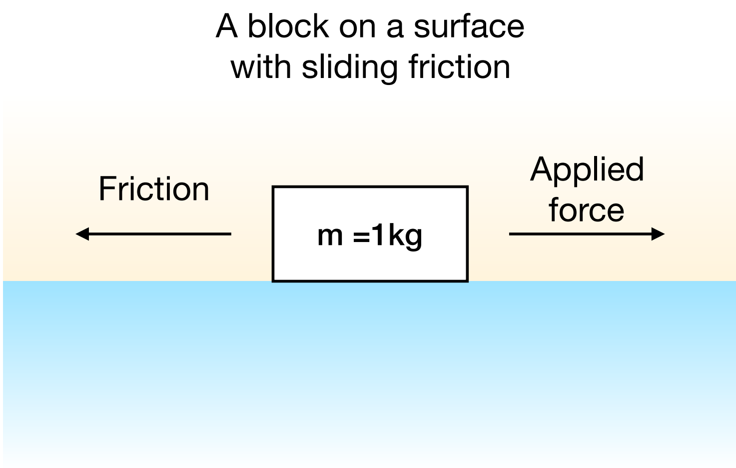


FIGURE 4.1: **The agent, a Bayesian cruise controller.** A block of mass = 1 kg, i.e., the agent, is placed on a surface with dynamic friction. The goal of the agent is to reach and maintain a velocity v_{des}.

tation sense) are added, cf. Chapter 6 where we also consider forces such as wind. To maintain consistency with the previous formulation, I still assume that w is weakly autocorrelated in a Stratonovich sense, i.e., not a Wiener process, even if the noise variables are implemented as white noise in my code[1]. The velocity measurement y is given as a (linear) noisy reading of x with observation noise $z \sim \mathcal{N}(0,\, \sigma_z^2)$:

$$y = x + z \tag{4.2}$$

The next step requires the definition of the agent's generative model, including a model of the system's dynamics:

$$x' = -\,\alpha x + v_{des} + w \tag{4.3}$$

[1]Wiener processes in continuous time are more often represented using a notation in terms of differentials dx, dt and dW where W is a Wiener process or Brownian motion. Some notable exceptions, however, use a Langevin formulation similar to the one applied here even in the case of white noise (Longtin, 2003; Longtin, 2010).

and of measurements:

$$
\begin{aligned}
y &= x + z \\
y' &= x' + z'
\end{aligned}
\tag{4.4}
$$

One of the major assumptions made in active inference is that the action variable a cannot be observed directly by an agent (i.e., it's not part of its generative model) and not necessary for problems of control (Friston et al., 2010a; Friston, 2011), giving rise to a different way of implementing homoestatic regulation. See in particular Chapter 7 for a thorough discussion. In active inference, one thus assumes that an agent is endowed with a minimal model that encodes how actions a modify observations y, y' (rather than hidden states x, x') via reflex arcs, as discussed in Friston et al. (2010a) and Friston (2011) and in all following chapters. In this case we also use, again for consistency, the notation in generalised coordinates of motion defined in Chapter 3 for random variables z, z'. Under Gaussian assumptions for z, z' and w, one can write the above state-space model in a probabilistic form:

$$
\begin{aligned}
p(y|x) &= \mathcal{N}(x,\, \sigma_z^2) \\
p(y'|x') &= \mathcal{N}(x',\, \sigma_{z'}^2) \\
p(x'|x, v; \alpha) &= \mathcal{N}(-\alpha x + v_{des},\, \sigma_w^2)
\end{aligned}
\tag{4.5}
$$

and considering the Laplace-encoded variational free energy defined in equation (3.30), here reported as

$$
F \approx -\ln p(\tilde{\psi}, \tilde{x}, \tilde{v}, \theta, \gamma)\Big|_{\tilde{\vartheta}=\tilde{\mu}}
$$

one can see that the probabilistic description of the generative model presented above reflects the likelihood and prior distributions necessary to build the generative density for the definition of the free energy in equation (3.30). Given equation (3.15) and specifying $\tilde{\psi} = \{y, y'\}$, $\tilde{x} = \{x, x'\}$, $\tilde{v} = \{v_{des}\}$, $\theta = \alpha$ and hyperparameters γ encoding properties about precisions $\pi_z, \pi_{z'}, \pi_w$, one gets

$$
p(\tilde{\psi}|\tilde{x}, \tilde{v}; \theta, \gamma) = \{p(y|x), p(y'|x')\}
\tag{4.6}
$$

and

$$
p(\tilde{x}, \tilde{v}; \theta, \gamma) = p(x'|x, v; \alpha)
\tag{4.7}
$$

The free energy then becomes (see equation (3.33)):

$$
F(y, \tilde{\mu}_x, \mu_v) \approx \frac{1}{2}\left[\pi_z (y - \mu_x)^2 + \pi_{z'} (y' - \mu'_x)^2 + \pi_w (\mu'_x + \alpha\mu_x - \mu_v)^2 - \ln(\pi_z \pi_{z'} \pi_w)\right]
\tag{4.8}
$$

with perception $\dot{\tilde{\mu}}_x = D\tilde{\mu} - \partial F / \partial \tilde{\mu}_x$, following equation (3.35), defined as:

$$\begin{aligned}
\dot{\mu}_x =& \mu'_x - \big[- \pi_z (y - \mu_x) + \pi_w \alpha (\mu'_x + \alpha \mu_x - \mu_v) \big] = \\
=& \mu'_x + \big[\pi_z (y - \mu_x) - \pi_w \alpha (\mu'_x + \alpha \mu_x - \mu_v) \big] \\
\dot{\mu}'_x =& \mu''_x - \big[\pi_w (\mu'_x + \alpha \mu_x - \mu_v) \big] = \\
=& - \pi_w (\mu'_x + \alpha \mu_x - \mu_v)
\end{aligned} \tag{4.9}$$

and action, $\dot{a} = -\partial F / \partial a$ from equation (3.36), as:

$$\dot{a} = -\frac{\partial F}{\partial a} = -\big[\pi_z (y - \mu_x) \frac{\partial y}{\partial a} + \pi_{z'} (y' - \mu'_x) \frac{\partial y'}{\partial a} \big] = -\big[\pi_{z'} (y' - \mu_{x'}) \big] \tag{4.10}$$

where we use the fact that an implicit model is embodied by the agent, with

$$\frac{\partial y'}{\partial a} = 1, \qquad \frac{\partial y}{\partial a} = 0 \tag{4.11}$$

see later chapters for more in depth discussions. These equations, when combined, form an action-perception loop with information inferred from the environment through perception and control exerted on the world via action. The combination of action and perception is regulated by precision parameters "π", representing weights in the weighted sum of prediction errors, see equation (4.8). Precisions encode the uncertainty (they are in fact inverse variances) of different variables of a generative model in an agent and effectively regulate the minimisation of free energy in equation (4.9) and equation (4.10). For the rest of this chapter and throughout this thesis, I will specify *sensory prediction errors* as the errors weighted by sensory precisions π_z, or more in general $\pi_{\tilde{z}}$, and *process* or *system prediction errors* as the ones weighted by process or system precisions π_w, or $\pi_{\tilde{w}}$, if dealing with generalised coordinates of motion. This distinction will be useful when I emphasise the role of precision weights on the minimisation of variational free energy, producing different behaviours based on their relative (sensory vs process, bottom-up vs top-down) strength.

More in general, precision parameters can be unrelated to the actual precisions of the true hidden states, causes and observations of a generative process (i.e., the world dynamics), and in some cases, as we will see, this misalignment becomes necessary (Feldman and Friston, 2010; Wiese, 2016). They have been addressed also in terms of "confidences", thought to encode how confident an agent is about its estimates of hidden variables. Precisions π's are in the most general case dynamic parameters that can change over time allowing for several types of behaviours to emerge depending on different situations, see for instance Feldman and Friston (2010) or Chapter 6 where these weights are optimised for different tasks. In this chapter, I assume fixed-valued precisions in order to focus, at least initially, on a few cases of "precision engineering" (Clark, 2015b) showing their role in the emergence of interesting and "pathological" behaviours (cf. Chapter 5). More specifically, I will analyse cases of "perception-centric" (or passive) and "action-oriented" (or active) agents within the context of active inference showing the importance of tuning precisions appropriately. Following Chapter 2, perception-centric agents heavily rely on perceptual inference, (over)focusing on estimating the causes of their sensory input, while action-oriented agents prioritise acting on the world over accounting for hidden properties of sensory input. After showing the importance of a closed sensorimotor loop, I will highlight some of the differences emerging in agents that focus on bottom-up (i.e., observations) or top-down (i.e., priors) information, using a precision weighting mechanism that emphasises different (i.e., sensory or system) prediction errors. Based on these ideas, I will present 4 case studies including agents having or not having access to action, and showing a dominance of bottom-up or top-down influences in their inference processes, as summarised in table 4.1.

TABLE 4.1: **The role of action in agents minimising free energy with different precisions' strengths.** The table summarises the results presented in terms of computational modes (purely bottom-up vs. purely top-down) and presence or lack of the ability to interact with the world (no action vs. with action).

	Bottom-up	**Top-down**
No action	Passive tracker	Passive dreamer
With action	Active tracker	Active dreamer

4.2.1 Just observing, the passive tracker

Passive trackers are agents that can only perceive their world without the ability of modifying any of its properties. They are an even more extreme version of the archetypical case advocated by CPP, already prioritising the estimation of the causes of observed sensations over adaptive behaviour. Passive trackers not only overprioritise perception over action, they also heavily rely on bottom-up observations over top-down priors, with precisions on sensory prediction errors (i.e., $\pi_{\tilde{z}}$, see equation (3.31)) taking a dominant role and driving predictions about incoming data. The larger the ratio between observation and system prediction errors, the smaller is the role played by prior beliefs. These agents present however in a straightforward way some of the arguments advocated by ideas of analysis by synthesis and the Bayesian brain hypothesis (Knill and Pouget, 2004; Yuille and Kersten, 2006), in particular the necessity of top-down information in the form of priors to disambiguate observations, whose estimates are otherwise entirely enslaved by bottom-up signals. As we can see in Fig. 4.2, in the simplest case, suitable priors filter out some of the noise in the measurements, separating the signal to be inferred (the black line) from the observation noise due to sensors/receptors.

If the precisions on sensory inputs are too large, there is no filtering, with increasingly more rugged predictions made just on the base of error correction processes

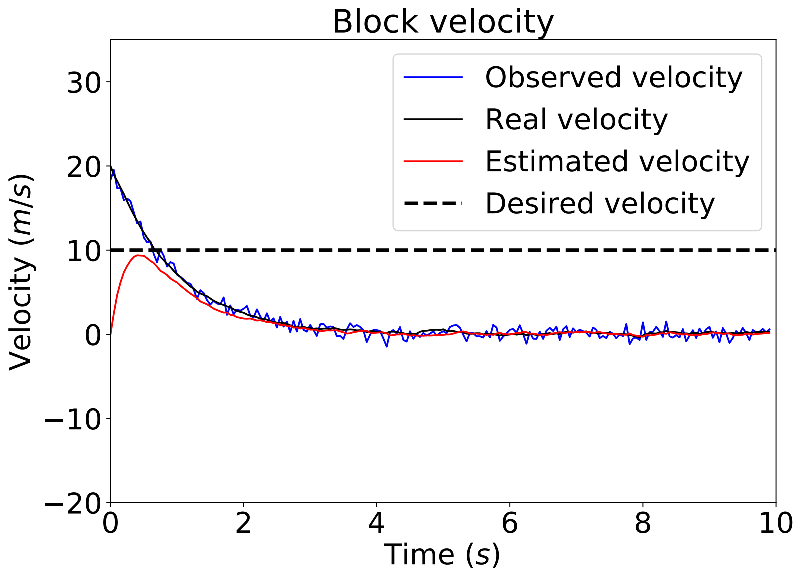


FIGURE 4.2: **(The passive tracker) The velocity of the block.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black).

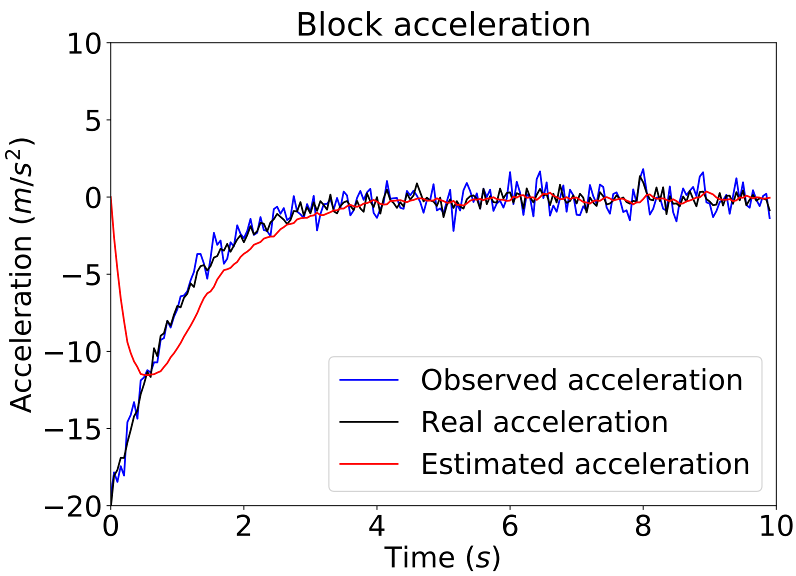


FIGURE 4.3: **(The passive tracker) The acceleration of the block.** The acceleration perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's real acceleration, i.e., without measurement noise (black).

triggered by new observations, see for instance Fig. 4.4. Given the Bayesian interpretation behind Kalman(-Bucy) filters (Meinhold and Singpurwalla, 1983; Chen, 2003) and upon which the FEP is built, this is consistent with results found in the Kalman filtering literature when the covariance matrices are off-balance and filters over-prioritise new measurements over previous estimates. On the other hand, one

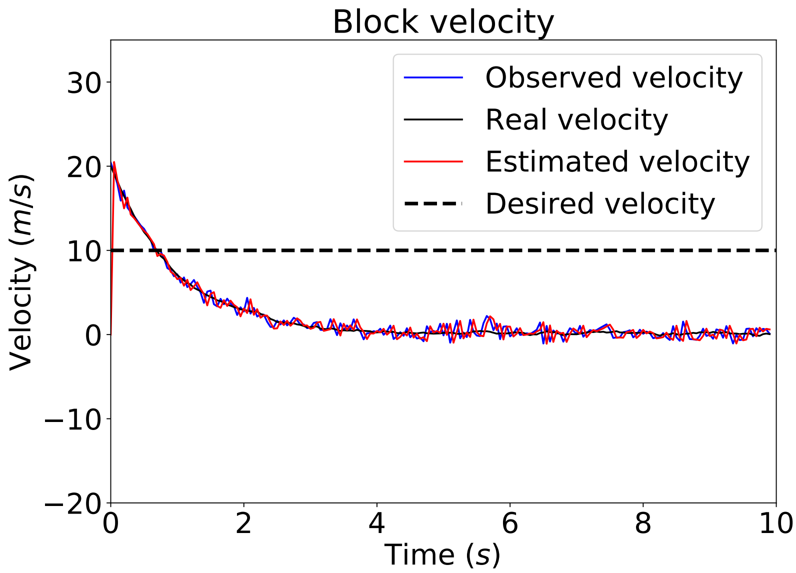


FIGURE 4.4: **(The passive tracker) The velocity of the block with higher sensory precisions.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black). In this case the sensory precision is increased to emphasise the tracking of the incoming signal.

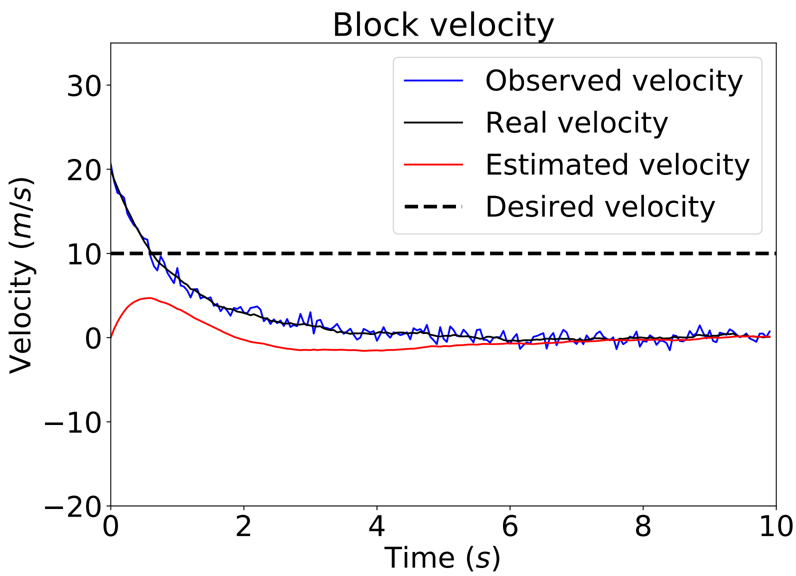


FIGURE 4.5: **(The passive tracker) The velocity of the block with lower sensory precisions.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black). In this case the sensory precision is decreased to emphasise the filtering properties of the recognition dynamics in equation (4.9).

can also show the complementary case, whereby slightly weaker sensory precisions allow for priors to filter more of the incoming sensory noise, showing less responsiveness to observation disturbances Fig. 4.5. This behaviour may be preferable in some cases, but at the same time it also induces slow responses to stimuli that may be relevant to the system (Sontag, 2003; Andrews, Yi, and Iglesias, 2006). See Chapter 6 for a more thorough discussion on this matter.

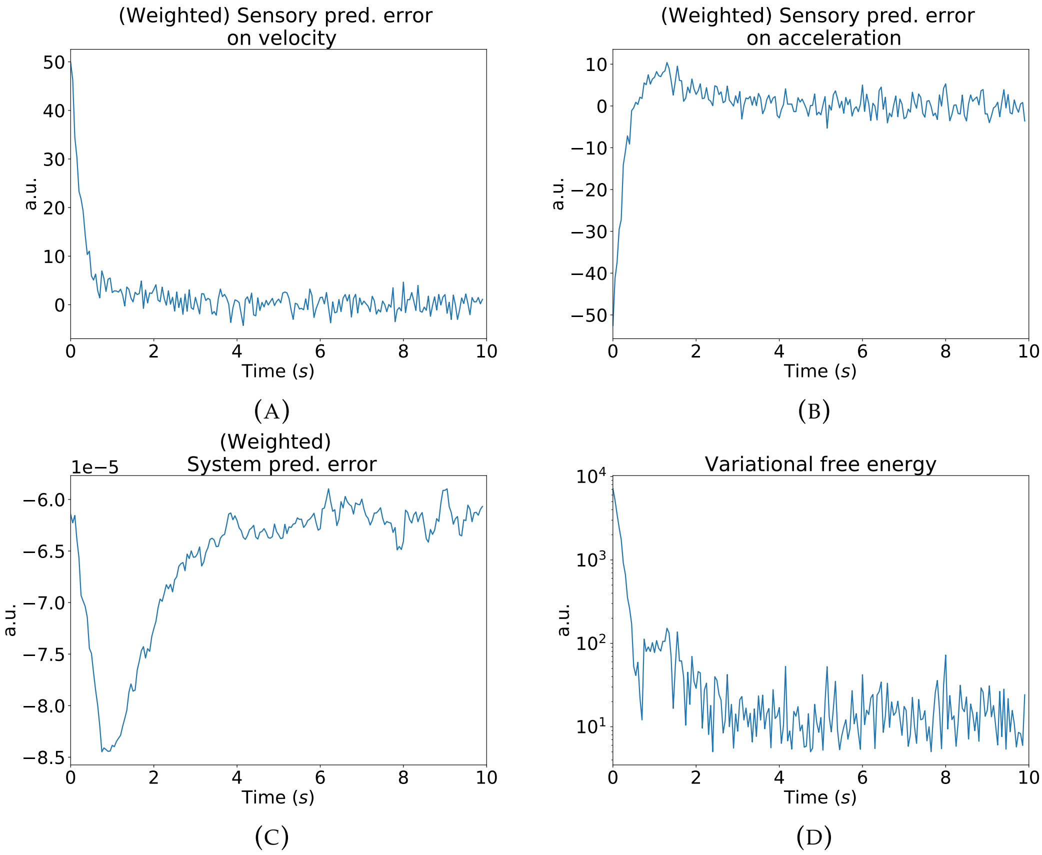


FIGURE 4.6: **(The passive tracker) Weighted prediction errors and variational free energy.** The evolution of sensory prediction errors on (A) velocity and (B) acceleration, and (C) system prediction error. (D) The variational free energy of the system over time.

In Fig. 4.6 we can see that the variational free energy of our initial agent (see Fig. 4.2) is (on average) minimised over time (Fig. 4.6d), driven mainly by the weighted prediction errors on sensory input, since sensory precisions $\pi_z, \pi_{z'}$ are much larger than process precision π_w. Both the weighted sensory prediction errors present a magnitude varying in the order of 10^1 (Fig. 4.6a and Fig. 4.6b), while the system error is in the order of 10^{-5} (Fig. 4.6c).

4.2.2 In a delusional state, the passive dreamer

Passive dreamers are also agents with no direct access to motor actions, as in the case of passive trackers. These agents prioritise their priors over incoming sensations. As we will see in an example later, this may become a feature when motor actions are introduced, but in the more basic case considered here, this simply implies the presence of agents with unfulfilled predictions about the world. These agents live in a sort of hallucinatory state where predictions are completely disconnected from sensory input, suggesting some possible conceptual link to models of psychiatric disorders (see for instance Friston (2005b) and Adams et al. (2013) and the discussion at the end of this chapter).

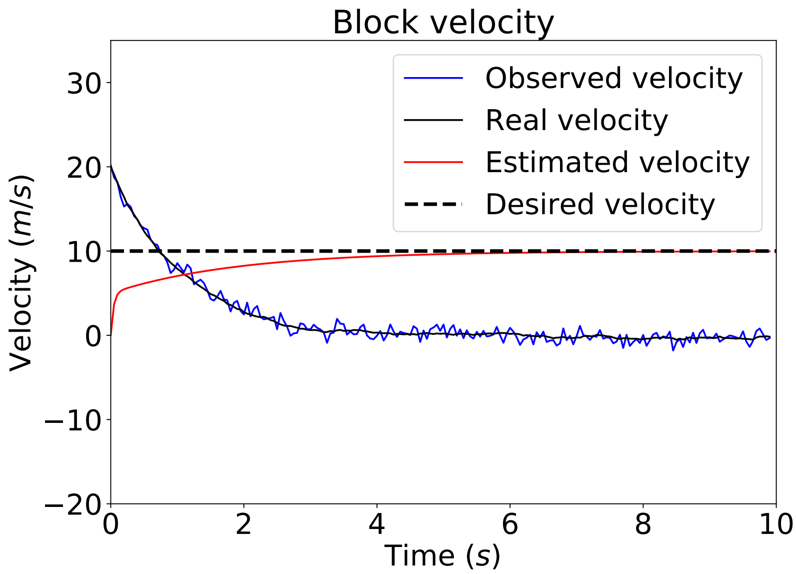


FIGURE 4.7: **(The passive dreamer) The velocity of the block.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black).

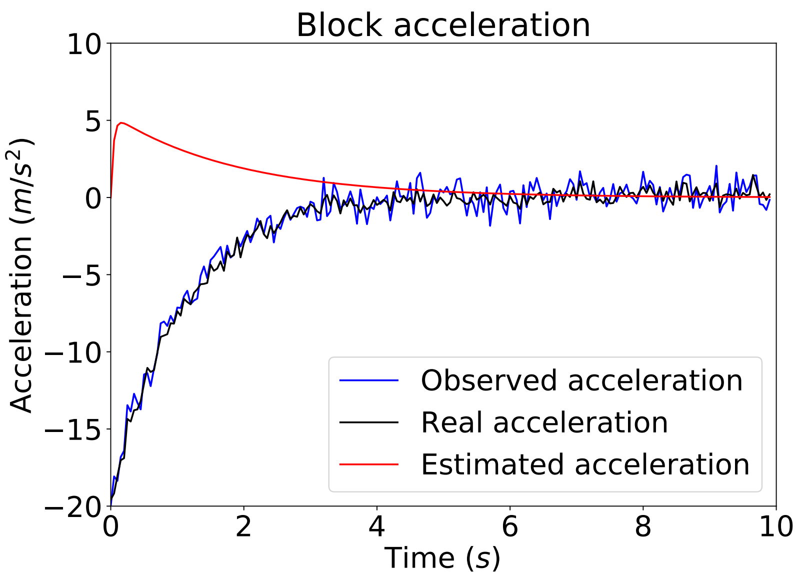
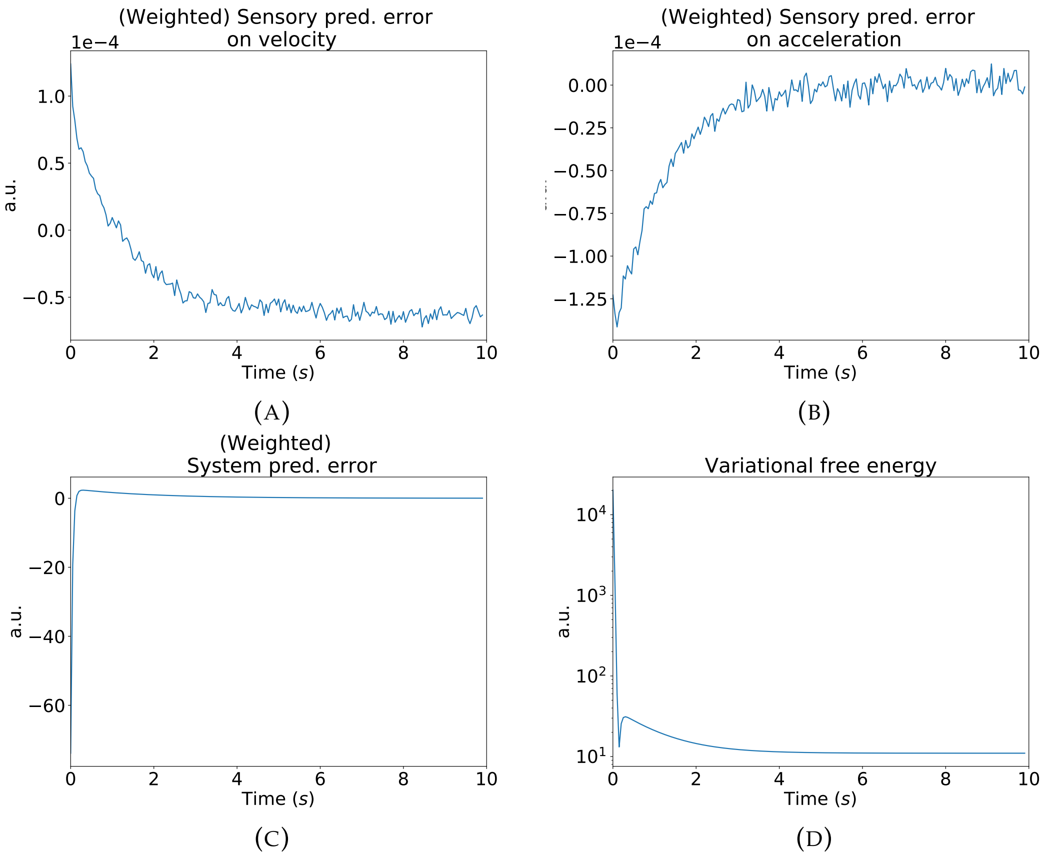


FIGURE 4.8: **(The passive dreamer) The acceleration of the block.** The acceleration perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's real acceleration, i.e., without measurement noise (black).

In passive dreamers, precisions on system errors π_w are overwhelming the ones on sensory prediction errors, driving the recognition dynamics for perception exemplified in equation (4.9) purely based on a model of the world that does not take into account new incoming information. For the velocity of the agent, we see a strong contrast between what the model describes, an equilibrium at $\mu_x = \mu_v = 10$ m/s when the system settles (red line), and the real velocity of the block which slowly decreases down to zero following the true dynamics of the world (black line), see Fig. 4.7. As for the acceleration, both the generative process and the generative model describe a system that eventually stops and since no force is applied (i.e., no external input and no action), both the estimated and the real accelerations reach the same equilibrium, see Fig. 4.8.



(A) (B) (C) (D)

FIGURE 4.9: **(The passive dreamer) Weighted prediction errors and variational free energy.** The evolution of sensory prediction errors on (A) velocity and (B) acceleration, and (C) system prediction error. (D) The variational free energy of the system over time.

As in the case of the passive tracker, the variational free energy is minimised (on average) over time, Fig. 4.9d, showing a sudden initial decrease due to the minimisation of the weighted prediction error on the system's dynamics and a temporary increase right after (Fig. 4.9c) due to the overshooting of the estimate of the acceleration (as observed in Fig. 4.8 and on a smaller scale in Fig. 4.9c). For the passive dreamer, the influence of sensory prediction errors is extremely low, with a difference of at least 5 orders of magnitude (Fig. 4.9a and Fig. 4.9b) when compared to the weighted prediction error on the system dynamics.

4.2.3 Acting with no reason, the active tracker

Active trackers are agents that can actively interact with their environment and unlike their passive version, they integrate action to close the sensorimotor loop. However, they are just another (although more elaborate) example of the perception centric description introduced by Clark, 2015a; Clark, 2015b, a direct consequence of Bayesian brain/predictive coding schemes (Rao and Ballard, 1999; Huang and Rao, 2011; Spratling, 2016) endowed with simple mechanisms for active behaviour and motor control. These agents can impact their environment through motor actions but they only do so to better sample sensations in agreement with their existing predictions, producing a "kind of self-fulfilling prophecy" (Hohwy, 2013; Clark, 2015a) entirely driven by incoming sensory input. Active trackers don't use (possibly relevant) priors to estimate their sensations and, as in the case of the passive tracker, are completely enslaved by their observations in a state of pure information gathering. Actions are only produced to cancel sensory prediction errors, to generate more accurate predictions about the world. Effectively, this creates the "dark room problem" for *active* agents exposed by Friston, Thornton, and Clark (2012), i.e., agents that "predict", or rather account for, all their observations, with actions simply bound to produce a process of inconclusive behaviour (unless the purpose for a system is to just estimate the hidden properties of its observations, unlike ours!).

The estimates of velocity, Fig. 4.10, and acceleration, Fig. 4.11, become good descriptions of the real variables in the world as in the case of the passive tracker. In the passive tracker example, however, the block naturally slowed down and eventually stopped (nearly stopped, because of the presence of noise) close to the origin. In the active version of the tracker, the initial sensory prediction error given by the estimate μ_x initialised at 0 triggers an action (see Fig. 4.12) which is then maintained constant over time after the prediction error on velocity is minimised, i.e., when the agent can predict its velocity. Having no other drive but to accurately predict its observations, this agent maintains its action constant since it has no cost (e.g., energy). Random initialisations of μ_x give different set-point equilibriums to the system, providing then simply different, but still accurate, estimates of the block's motion.

As in the case of the passive tracker, sensory prediction errors (Fig. 4.13a and Fig. 4.13b) exert a much larger influence on the minimisation of variational free energy (Fig. 4.13d) due to the precision weighting mechanism enforcing their role. The only significant difference between the active and the passive versions is on the process prediction error, cf. Fig. 4.13c and Fig. 4.6c, given by the fact that the active tracker gets further away from the "desired" state represented by the prior thanks to its motor actions.

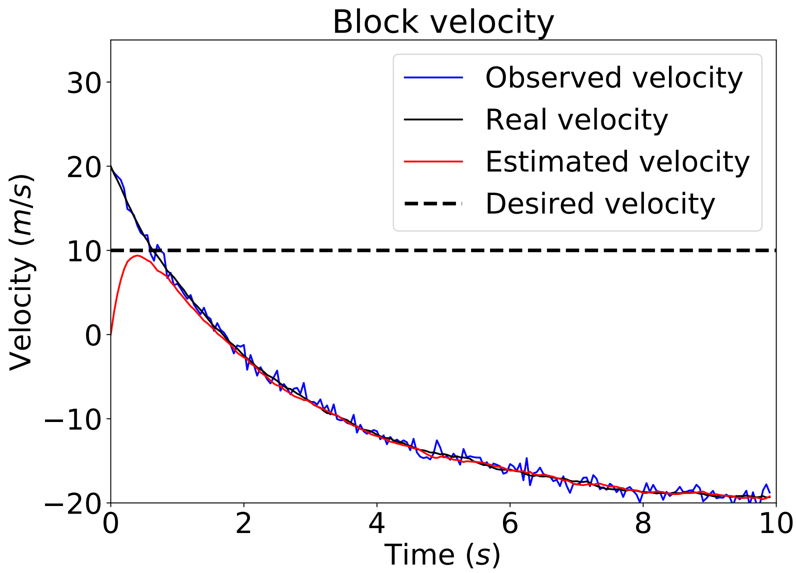


FIGURE 4.10: **(The active tracker) The velocity of the block.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black).

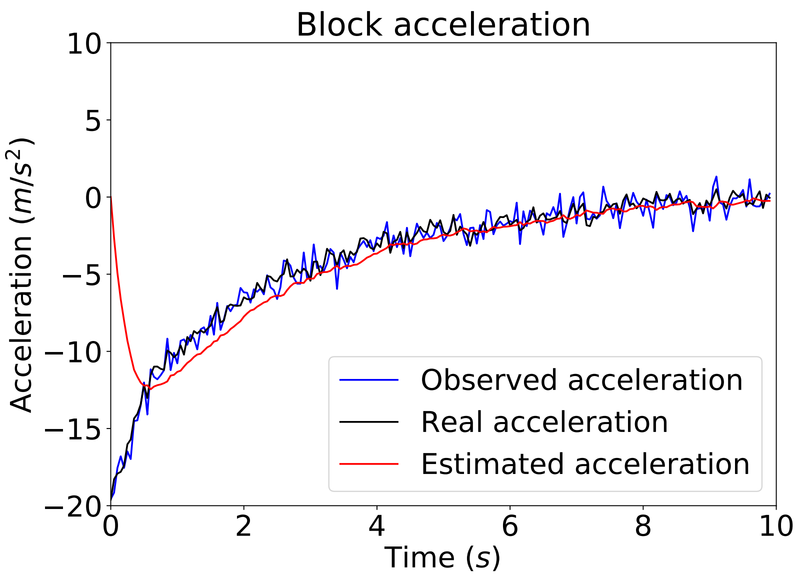


FIGURE 4.11: **(The active tracker) The acceleration of the block.** The acceleration perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's real acceleration, i.e., without measurement noise (black).

4.2.4 Chasing one's dreams, the active dreamer

The active dreamer presents a similar delusional state to the one described by its passive counterpart. Unlike the passive dreamer, however, the active agent has the

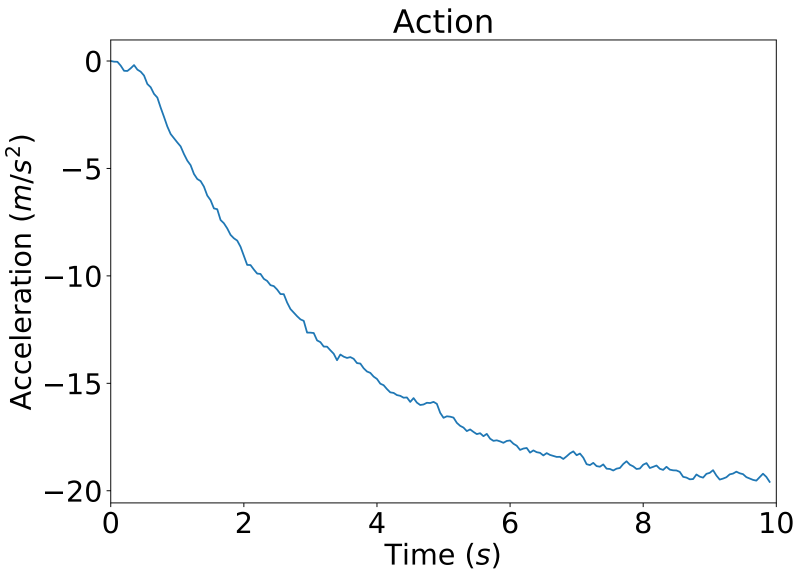


FIGURE 4.12: **(The active tracker) The motor action of the agent.** The action induced by the minimisation of variational free energy following active inference given, in this case, a weak prior.

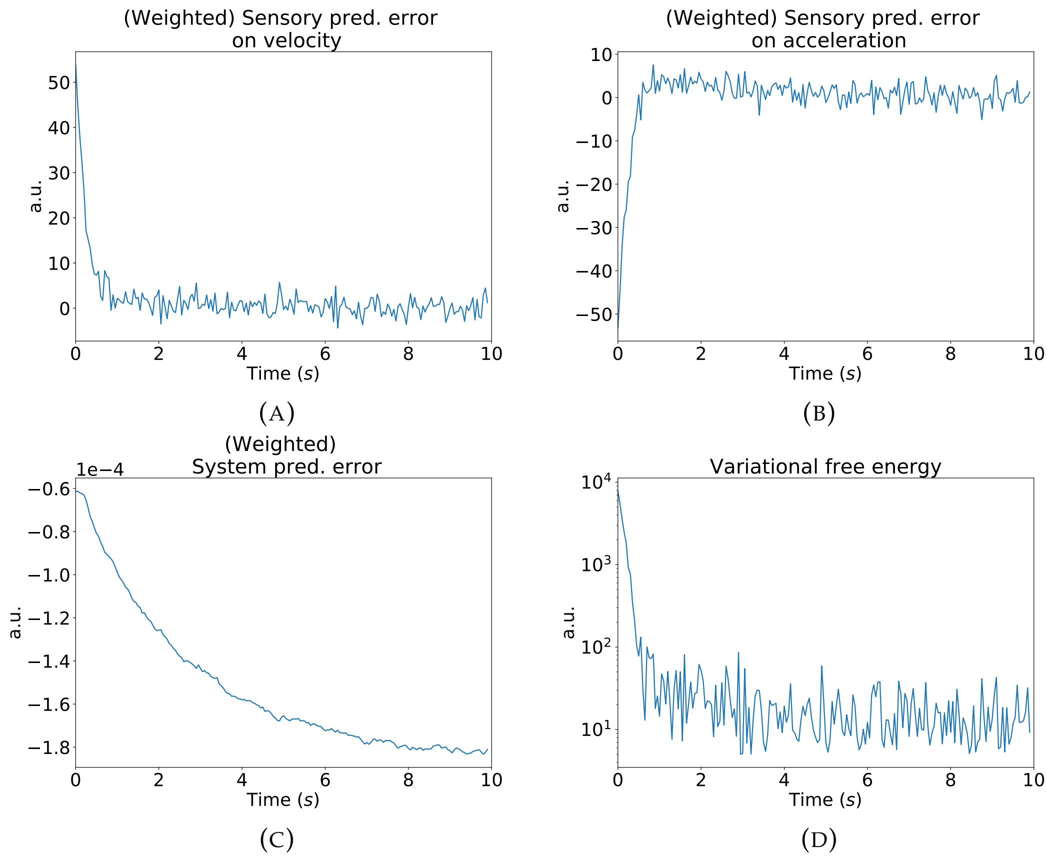


FIGURE 4.13: **(The active tracker) Weighted prediction errors and variational free energy.** The evolution of sensory prediction errors on (A) velocity and (B) acceleration, and (C) system prediction error. (D) The variational free energy of the system over time.

"power" to realise the desires encoded by its priors. This hallucinatory state becomes then a feature in active inference, following the equivalence between (stochastic) optimal control and Bayesian inference (Todorov, 2009b), recasting the problem of regulation as a problem of inferring the necessary actions to fulfil observations biased by the agent's priors. In this picture, self-generated sensory prediction errors trigger actions that bring about the desired state of an agent. Unlike the active tracker, actions here serve a purpose encoded by the agent's priors and while this may not be satisfactory by itself (i.e., where do these priors come from?), it provides a possible description of a normative account of behaviour once these priors are appropriate and serve an agent's needs.

As we can see in Fig. 4.14, the estimate of the velocity is initially rather poor, given the strong role played by priors that do not align with the world dynamics.

This misalignment, however, is what triggers action in Fig. 4.16, since it's the only other means to minimise variational free energy. Both the estimates of velocity (Fig. 4.14) and acceleration (Fig. 4.15) become accurate descriptions of the true readings once the desires are realised by action, creating the "self-fulfilling prophecy" advocated by Hohwy (2013) and Clark (2015a). In this case, however, the "prophecy" is realised through adaptive behaviour following an agent's drives. The variational free energy landscape and its minimisation are then entirely consistent with the passive dreamer case. The only substantial difference is in the fact that sensory prediction errors are now more effectively minimised, especially the one on velocity Fig. 4.17a, thanks to the presence of motor actions affecting the agent's observations now increasingly more in agreement with its "delusional" predictions.

TABLE 4.2: **Agents' parameters and setups.** The table summarises the parameters used to simulate our two agents, the passive tracker and the active tracker, following the implementation of equation (4.9) and equation (4.10).

	π_z	$\pi_{z'}$	π_w	Action
Passive tracker	$\exp(1)$	$\exp(1)$	$\exp(-12)$	$a = \dot{a} = 0$
Passive dreamer	$\exp(-12)$	$\exp(-12)$	$\exp(2)$	$a = \dot{a} = 0$
Active tracker	$\exp(1)$	$\exp(1)$	$\exp(-12)$	$\dot{a} = \partial F/\partial a$
Active dreamer	$\exp(-12)$	$\exp(-12)$	$\exp(2)$	$\dot{a} = \partial F/\partial a$

4.3 Discussion

In theories derived from the Bayesian brain hypothesis (Knill and Pouget, 2004; Doya, 2007) and predictive processing (Hohwy, 2013; Clark, 2015b), there is often a strong emphasis on perceptual processes. This is both due to historical reasons that trace these ideas back to work by Helmholtz, and related, modern theories of analysis by synthesis (Helmholtz, 1867; Neisser, 1967; Gregory, 1970), and to a strong

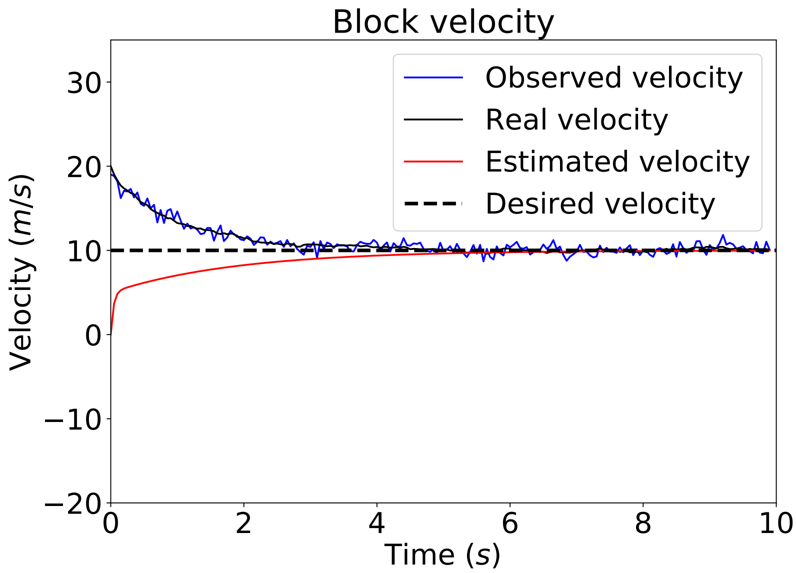


FIGURE 4.14: **(The active dreamer) The velocity of the block.** The velocity perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's true velocity, i.e., without measurement noise (black).

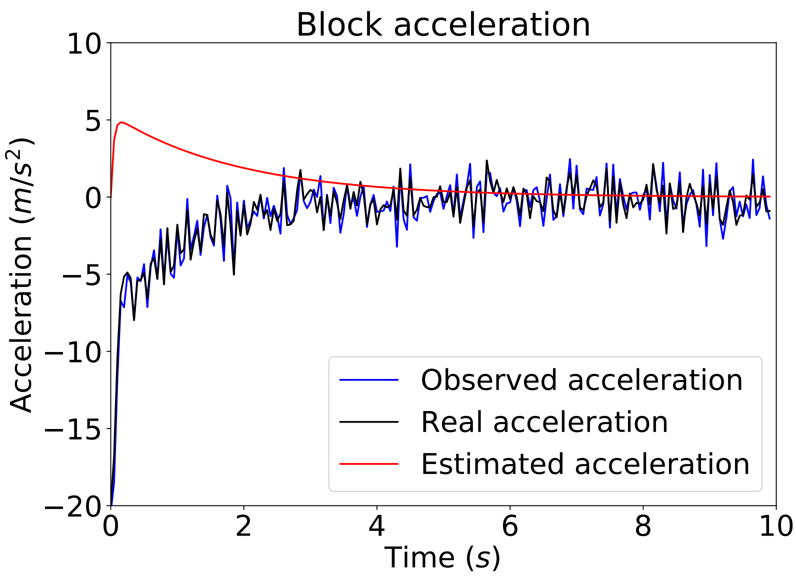


FIGURE 4.15: **(The active dreamer) The acceleration of the block.** The acceleration perceived by the agent (blue line), its best estimate according to weak priors (red) and the block's real acceleration, i.e., without measurement noise (black).

tradition in the cognitive sciences to focus on perception and cognition over action and behaviour (Fodor, 1983; Boden, 2006). The repercussions of this bias in Bayesian theories of the mind are deep and rooted, constantly re-emerging even in the most modern proposals on the Bayesian brain.

Following the definition given by Clark (2015a) and Clark (2015b) and presented

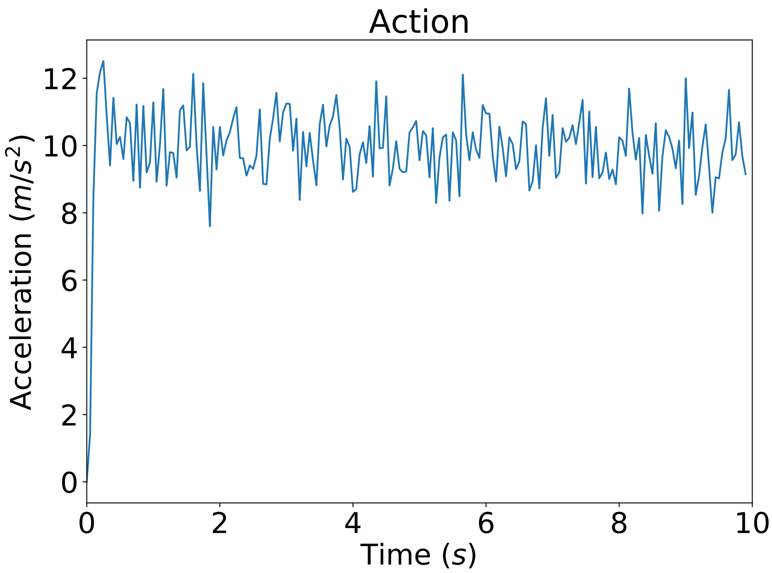


FIGURE 4.16: **(The active dreamer) The motor action of the agent.** The action induced by the minimisation of variational free energy following active inference given, in this case, a strong prior.

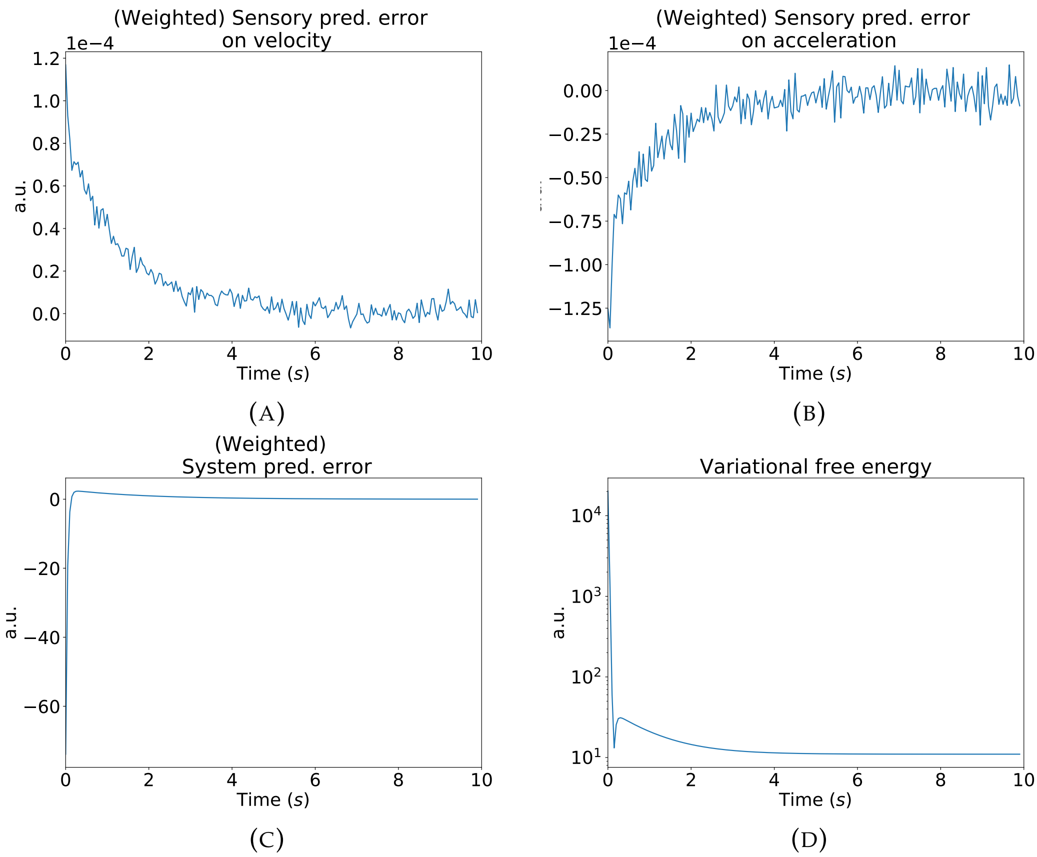


FIGURE 4.17: **(The active dreamer) Weighted prediction errors and variational free energy.** The evolution of sensory prediction errors on (A) velocity and (B) acceleration, and (C) system prediction error. (D) The variational free energy of the system over time.

in Chapter 2, I also strongly advocate for a formal distinction between "perception-centric" and "action-oriented" Bayesian approaches to cognitive science (see also Engel, Friston, and Kragic (2016) and the idea of passive and active agents in Chapter 1), with implications that could potentially capture aspects of the more general discussion between computationalist and 4E approaches to cognitive science.

In this chapter I provided a simple model of a sensorimotor loop built using active inference and aimed at showing, with an example of homeostatic regulation, some of the most fundamental points and possible misunderstandings of the FEP and related theories. It was my goal to make the example especially simple, and for this reason the problem of regulation was reduced to a (Bayesian) cruise controller for an agent (i.e., a block) sliding on a surface with dynamic friction. The friction naturally slows the block down, but the agent is endowed with ability to apply a force over time that allows the block to move and maintain a desired speed. With this example I then explored a few cases, 4 in total, combining the crucial components of an active inference regulator: a closed action-perception loop and a specific balance of weighted of prediction errors to minimise variational free energy according to the goals of the agent. During this exploration I introduced two agents which I named the "passive tracker" and the "passive dreamer", where actions on the world were not allowed. The weaknesses of stories without motor actions became soon obvious, but it was nonetheless important to establish the background over which the rest of the simulations in this chapter, and more in general in this thesis, is based (see also Bruineberg, Kiverstein, and Rietveld (2018), where this point is explored considering the literature on Bayesian brain and predictive coding theories). Alongside the absence or presence of action to define an appropriate sensorimotor loop, I also explored the balance of different (weighted) prediction errors. As we can see in equation (4.8) in fact, the expression for the variational free energy under the Laplace approximation (see Chapter 2) is reduced almost entirely to a weighted sum of prediction errors. These prediction errors can be divided into sensory and process (or system) prediction errors, the former encoding mismatches between predictions and (bottom-up) sensory data, the latter encoding the differences between (top-down) prior information and the current best estimate of the hidden variables of a system. All these prediction errors are weighted by precision hyperparameters, encoding the inverse variance of observations and hidden dynamics of a system. As stressed in previous chapters and shown more in detail here, these hyperparameters need not encode true properties of the world and can instead be seen as quantifying the uncertainty, or confidence, of an agent in its estimates. Considering that precisions are, in principle, defined over a continuous interval of values, I simplified the analysis by imposing either higher sensory $\pi_{\tilde{z}}$ or high process precisions π_w. High precisions drive the minimisation of free energy, enforcing the relative strength of one subset of hyperparameters and relative prediction errors over the other.

The two examples where no action was included, the passive tracker and the passive dreamer, represent two extreme versions of purely bottom-up and top-down driven perceptual processing. The former passively engages with the world, attempting to estimate new observations. The complete lack of prior information, however, forces this agent to rely entirely on new observations and so, at best, to track the incoming sensations over time after they have been observed. For the passive tracker, every sensation is essentially "surprising" (in statistical terms) since priors play little to no role in making predictions about incoming data. Sensory prediction errors have a much larger amplitude and are thus driving the minimisation of free energy. The passive dreamer on the other hand, doesn't take into account newly observed sensations and is stuck in constant delusional state imposed by its strong priors, e.g., the velocity being 10 m/s. The state of affairs of the world is, in this case, completely insignificant since the agent simply hallucinates its desired outcomes. Only process prediction errors are minimised, since sensory precisions are too low to play a role in the optimisation of variational free energy. In this set up, none of the new sensations are "surprising" (again, in statistical terms) since the agent is completely uninterested in new incoming observations. Perhaps at odds with our common intuition, this agent however presents some of the crucial elements necessary for the active inference version of a regulator. The self-generated misalignment of real sensations and predictions about them biased by the agent's desires constitutes in fact the basis for forming (proprioceptive) prediction errors that can be resolved using action (Friston et al., 2010a; Brown et al., 2013; Wiese, 2016).

In this context, it is easy to see why homeostatic regulation requires both a perceptual process of estimation of the world (i.e., the agent's velocity) and in particular, an action selection procedure based on a "desire", i.e., an internal set-point/target encoded in the form of a prior. To highlight, once again, the role of prior information for an agent I proposed two different agent in this case too: the "active tracker" and the "active dreamer". The active tracker largely follows the fate of the passive one, bound to simply attempt to account for its observations. Action, in this case, simply speeds up the process and generates behaviour similar to the one described by the dark room problem (Friston, Thornton, and Clark, 2012). An agent whose only purpose is to predict its sensations should find a state/place where sensations are trivially predicted, i.e., a dark room. Considering the block in our set up, the closest state to a "dark room" is any equilibrium of the system reached when the action is stationary, since the action has no associated cost. This agent simply finds the best way to predict its state by bending the world to its predictions and generating predictions that better conform to its sensations, no matter what the state of the environment is. It is somewhat similar to the idea of niche construction in evolutionary biology (Odling-Smee, Laland, and Feldman, 2003) and considered in active inference too (Bruineberg et al., 2018; Constant et al., 2018), but lacks the idea of normativity constraints for building an environment that better promotes the survival of an organism.

The last agent presented is the active dreamer. This agent actively engages with the environment to change the incoming sensations to match its priors, combining top-down and bottom-up processes with a more balanced role for precisions, still however emphasising top-down information. Unlike the passive tracker and dreamer, this agent interacts with the world and can introduce real changes to its perceived input. Unlike the active tracker that predicts its sensations by just performing actions that can fulfil them, our agent regulates variables in the environment in order to generate input more in line with its priors. As previously mentioned, the passive version of this agent may have looked a bit puzzling: strongly predicting a state of affairs that has no correlation to states in the world while discarding all the new information may look unintuitive. However, once action is introduced, this "misperception" of the environment becomes extremely valuable. As we will see in detail in Chapter 7, standard approaches to problems of control used in biology and neuroscience are still based on accurate inferences of world variables, building motor control out of inverse models mapping accurate estimates of unobservable states of the environment to an agent's motor commands. While this is (in principle) possible in the linear case, the complexity of any realistic implementation of inverse models in both biological and artificial systems is prohibitive (Adams, Shipp, and Friston, 2013; Friston, 2011). In active inference, this process is bypassed by turning the control problem into an inference one (see also Todorov (2009b) for a clear explanation), having agents that create proprioceptive prediction errors that can only be resolved by acting on the world (Friston, 2011). In our example, the proprioceptive modality is represented by readings of the agent's velocity and acceleration, and errors are created by imposing a prior on the expected motion of the agent: if the agent expected to maintain a velocity of 10 m/s, what action would it have to take to achieve such goal? Fixing the final objective and calculating (trivially, in this case) the time-independent policy by backward induction, one then finds a solution given by a set of actions that counteracts the effects of sliding friction (responsible for slowing the block down).

Even on such a simple system, some further investigations may be possible but will not be tackled now. In particular, a possible directions involves the exploration of the possible connections to computational psychiatry (Montague et al., 2012; Stephan and Mathys, 2014). The nascent field of computational psychiatry seeks to apply knowledge derived from computational models of decision making to psychiatric nosology (i.e., the classification of diseases) in the standard diagnosis process (Redish, 2004; Montague et al., 2012; Huys, Maia, and Frank, 2016; Redish and Gordon, 2016). The underlying idea is that aberrant behaviours found in psychiatric disorders can be attributed to anomalous functioning of the (computational) mechanisms responsible for the generation of such behaviours and other atypical experiences related to a misperception of the world. Classifying pathologies in patients based on behaviour alone remains pervasive in modern diagnosis procedures but the field is decisively moving in a new direction. The goal of computational psychiatry is to pursue an understanding of behaviour and decision making based on the study of the latent (computational) mechanisms of psychiatric disorders through the use of mathematical and computational models, in order to improve treatments of different conditions. In a proposal I wrote for a fellowship in 2017, I argued for the use of this approach not only for human experiments where it is in principle applicable (e.g., Redish (2004)), but also on artificial systems with the aim of 1) testing the proposed computational techniques on systems with no side effects, i.e., artificial ones, and 2) improve the mathematical analysis and study of solutions that in modern machine learning and AI tend to prefer "black-box" approaches. With regards to this chapter's agents, one could imagine to study the passive tracker's state as a very simple and high-level (phenomenological) model of schizophrenic traits (Adams et al., 2013; Friston et al., 2016b). By building appropriate and more realistic mechanisms of precisions update (see Chapter 6) or state-dependent uncertainty (see Feldman and Friston (2010)), one could simulate the inability of modulating the precision of sensory errors relative to the priors, studying how this produces a series of effects (e.g., resistance to illusions) comparable to signs and symptoms of schizophrenia and other psychiatric conditions. One of the main risks, in this case, would be to oversimplify the analysis of mental health disorders by producing models that are extremely detached from reality or that make use of biologically or behaviourally unreasonable assumptions. This is also the case, more in general, with Bayesian descriptions of any phenomenon (Pearl, 2001; Jones and Love, 2011; Bowers and Davis, 2012): accepting any prior as an explanation without having to commit to the "right" answer. This discussion will however be expanded on in the final Conclusions, where I will focus on some of the potential weaknesses of the FEP.

Another example of overly generous interpretations allowed by the use of active inference comes from the interpretation one can give to precisions. As inverse variances of variables in the generative model, they represent the uncertainty of an agent on different variables of a model. In equation (4.8), as explained before, they implement a weighting mechanism for different prediction errors. This is then used in equation (4.9) and equation (4.10), where one can also see that their role as weights becomes now similar to "learning rates" in connectionists' models of optimisation. In most of the cases reported in this chapter, these different interpretations are simply complementary and shed light on different aspects of the role of precisions. In the last simulation implementing the active dreamer, however, I introduced an extra learning rate in equation (4.10) to allow the convergence of action in a reasonable time. While this may raise questions on the meaning of this extra parameter, its presence in my simulations is only due to the method used for solving stochastic differential equations in my examples, the Euler-Maruyama algorithm. This algorithm is extremely straightforward, but it can generate large errors for even simple stochastic simulations, as in my case when I tried to use very large process precisions in the active dreamer. To avoid such issues, I would have to implement simulations with a step size $dt < 10^{-8}$ which would make most of my code too computationally demanding even in its simplest form. In similar examples, see for instance Friston, Trujillo-Barreto, and Daunizeau (2008), Friston uses a local linearisation approach (Ozaki, 1992) which recasts a continuous time SDE into a discrete one while preserving most of its properties, and then simulates an exact discrete-time process.

In my work I also claimed that only the relative dominance of one set of precisions over the other defines which prediction errors have priority in the minimisation of variational free energy. More in general, however, work such as Feldman and Friston (2010) makes explicit and different predictions on the behaviour of a system when sensory precisions are attenuated or process ones are strengthened. According to this hypothesis, the former defines models of the common (and according to the authors necessary) effect of sensory attenuation, which would enable agents to effectively generate motor action under active inference assumptions that, as we saw, require not "paying attention" to our actual sensations. The latter, instead, is proposed to be responsible for hallucinatory-like states where overly eager organisms imagine different sensory inputs influenced by priors that are over-influencing estimates of the state of the world.

4.4 Conclusion

This chapter presented a minimal implementation of active inference for the definition of an action-perception loop. The example was intentionally simple in order to show the basic elements of an active inference agent and to focus on two aspects of this framework: 1) the importance of action, to highlight the limitations of accounts of the Bayesian brain hypothesis and predictive processing relying almost entirely on models of perceptual inference, and 2) the role played by a balanced set of precisions, to emphasise that simple estimations of latent variables in the world with weak priors does not allow the implementation of normative behaviour in an agent (unless the norm is to just estimate hidden variables). The first point may look obvious, but it's often overlooked by implementations of predictive processing/coding models focusing on models of perception, implying that behaviour will emerge as a consequence of good models of the world generated purely via perceptual accounts of incoming sensory input. The second one builds on this idea and extends it to show one of the crucial differences of the active inference framework: behaviour can only be created by inaccurate inferences of the world. Unlike other similar frameworks based on control theory requiring accurate estimates of latent variables (see Chapter 7), in active inference motor actions and behaviour can emerge only in order to minimise prediction errors created by systematic misrepresentations (Wiese, 2016).

Chapter 5

Generative models of sensorimotor contingencies

Previous implementations of active inference have typically relied on a passive, “perception-centric" view of this theory, assuming that agents are endowed with a detailed generative model of their surrounding environment. These “perception-centric" approaches subordinate motor actions to the accurate and comprehensive perception of the environmental properties generating sensory data (Hohwy, 2013) (cf. the classical sandwich of cognitive science (Hurley, 2001)) and thus have often brought the FEP and active inference into direct conflict with 4E views of cognition. In contrast to these claims, we present here an “action-oriented" (Clark, 2015a) solution showing how adaptive behaviour can emerge even when agents operate with a simple model which bears little resemblance to their environment, focusing on a more ecological and embodied reading of the FEP (Seth, 2014b; Clark, 2015a; Bruineberg, Kiverstein, and Rietveld, 2018; Allen and Friston, 2018). This view is, as we will see, more in line with 4E theories of cognition highlighting that behaviour generated by apparently complex world models is, in fact, the product of agent/environment dynamical couplings. In this light, agents can only be properly studied and understood when coupled to their environment and generative models are, essentially, not generating objectively good predictions of the world but simple sensorimotor contingencies.

In this chapter we implement an example of a simple wheeled agent performing phototaxis under active inference and present it as a proof of principle of an “action-oriented" reading on the FEP. We will explain the assumptions that allow for the emergence of phototaxis, showing its dependence on precision parameters (inverse variances) of different prediction errors and finally present an example of a different, “pathological" behaviour when some of these assumptions are not met. This chapter is based on Baltieri and Buckley (2017), with a few improvements: the state-space model is now represented correctly using random variables rather than their means (cf. Buckley et al. (2017)) and with learning rates (still introduced here for completeness) now all set to 1, simplifying our interpretation of the agent’s behaviour which is now completely independent on extra parameters. Compared to Chapter 4, here I strongly emphasise the differences between the generative model of the agent and the generative process of the world (i.e., what characterises action oriented agents in active inference), closer to ideas of sensorimotor contingencies (Buhrmann, Di Paolo, and Barandiaran, 2013; Seth, 2014b) rather than agents as mirrors of the world.

5.1 Background

As shown in Chapter 3, under a series of assumptions (mainly the Laplace approximation), the free energy functional simplifies to equation (3.30). For simplicity, here we assume that no hidden inputs $\tilde{v}$ are represented in the variational free energy. Furthermore, in this work we do not consider the use of generalised coordinates of motion, simplifying thus variables $\tilde{\psi}, \tilde{x}$ to ψ, x:

$$F \approx -\ln p(\psi, x; \theta, \gamma)\Big|_{\vartheta=\mu} \tag{5.1}$$

Parameters θ are all equal to 1 while hyperparameters γ are fixed variables; both of them are dropped in this formulation from now on and only discussed later on

$$F \approx -\ln p(\psi, x)\Big|_{x=\mu_x} \tag{5.2}$$

As seen in previous chapters, $p(\psi, x) = p(\psi|x)p(x)$ is the generative density comprising of a likelihood $p(\psi|x)$ and a prior $p(x)$ on (estimated) hidden states x. In this chapter we will consider multiple *independent* variables, thus maintaining a formulation in terms of scalar variables and function(al)s, and only in Chapter 7 present a multivariate implementation of the FEP. Under this framework, it is suggested that perception is implemented as the minimisation of variational free energy with respect to the means of hidden states μ_x. With assumptions on the generative model making external noise and internal fluctuations Gaussian random variables and under the Laplace approximation (MacKay, 2003; Friston et al., 2007) (see also Chapter 3), one can in fact reduce the optimisation of hidden states x to the minimisation of their modes/means μ_x, showing that the variances of the recognition density, ς^2, are recovered analytically with the Hessian of the variational energy (MacKay, 2003; Friston et al., 2007) (see again Chapter 3). Perception is thus implemented, as seen in Chapter 3, as:

$$\dot{\mu}_x = -\frac{\partial F}{\partial \mu_x} \tag{5.3}$$

which updates μ_x and converges when the minimum of the variational free energy F is reached, i.e., when $\partial F/\partial \mu_x = 0$. This equation, unlike equation (3.35) in Chapter 3, does not consider generalised coordinates of motion, meaning that higher embedding orders of motion of expected hidden states x are suppressed. In practice this means that our generative model will be in a static frame of reference (i.e., not chancing over time), and that measurement noise and internal fluctuations will be modelled as Gaussian random variables with zero autocorrelation in the more traditional Ito's sense, as in classic Bayesian filtering methodologies based on standard state-space model formulations (Friston, Trujillo-Barreto, and Daunizeau, 2008).

In contrasts to perception, action is defined as a process of changing the world such that sensory data better accords with predictions of the generative model (Fig. 5.1). Specifically, in terms of the formalism presented in Chapter 3, while perception minimises the divergence term in the definition of free energy equation (3.6) by finding a variational density closer to the posterior, action optimises indirectly the second one, surprisal, by updating sensations ψ (Bruineberg, Kiverstein, and Rietveld, 2018). To achieve this, an agent must know (or at least have an approximation of) how inputs ψ depend on motor actions a (i.e., $\psi = f(a)$) (Friston et al., 2010a; Buckley et al., 2017). With such knowledge, action can similarly be cast as a minimisation via gradient descent on the free energy with respect to a:

$$\dot{a} = -\frac{\partial F}{\partial a} = -\frac{\partial F}{\partial \psi}\frac{\partial \psi}{\partial a} \tag{5.4}$$

Thus action and perception can be described as the minimisation of the same quantity, with the simultaneous implementation of both processes closing the action-perception loop, see Fig. 5.1.

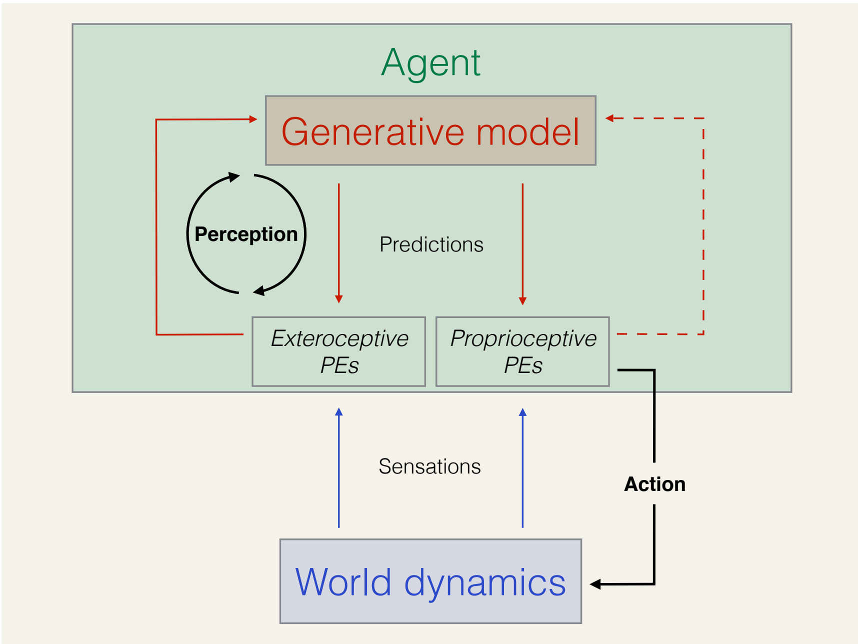


FIGURE 5.1: **A schematic of the FEP.** Two types of sensations, exteroceptive (e.g., light intensity, ψ_l) and proprioceptive (e.g., motor velocity, ψ_m), represent the sensory input of an agent (blue arrows). Beliefs or predictions on exteroceptive input are updated within the generative model through perception (red arrows). The dashed red arrow denotes the lack of updates on proprioceptive prediction errors due to new sensations, necessary for phototaxis later on. This update is however introduced later on to show "pathological behaviour". Action solves the discrepancy between predictions of the generative model and incoming sensations by engaging with the world.

5.2 A minimal generative model of phototaxis

To present some of the core ideas behind the FEP and active inference and their connections to 4E theories, we implement phototaxis on a simple 2-wheeled vehicle[1]. We simulate an agent with circular body, 2 noisy light sensors and 2 noiseless motors, see Fig. 5.2. For simplicity, we do not simulate occlusion of the light source by the agent's body.

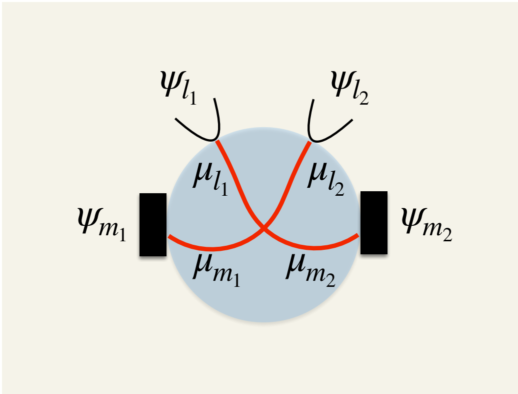


FIGURE 5.2: **The wheeled vehicle used in our simulations of phototaxis.** The agent receives input from two exteroceptors reading light intensity (ψ_{l_1}, ψ_{l_2}) and two proprioceptors reading wheel velocity (ψ_{m_1}, ψ_{m_2}). Variables $\mu_{l_1}, \mu_{l_2}, \mu_{m_1}, \mu_{m_2}$ are part of the generative model of the agent. The red lines represent the relations between the agent's priors on the dynamics of the world, very distant from how the real dynamics work.

In several previous agent-based simulations of the FEP, it is typically assumed that an agent possesses a rich and detailed model of its environment, matching the complexity of the dynamics describing the world, see for example Friston and Kiebel (2009a), Friston et al. (2010a), and Friston et al. (2012). If we were to take the same approach here, we would perhaps start by assuming that the agent has a representation of the locations of both itself and the light source, following a methodology similar to SLAM (Simultaneous Localisation and Mapping) (Thrun, Burgard, and Fox, 2005). However, a more action-oriented interpretation of the FEP suggests that adaptive behaviour could emerge from generative models that are more parsimonious and encoding only sensorimotor contingencies (Seth, 2014b; Clark, 2015a), capturing only basic regularities of the coupled agent/environment system. To examine this, we endow our agent with a minimal model of its surrounding environment. Specifically, our agent receives four inputs: two from exteroceptors sensitive to light ψ_{l_1}, ψ_{l_2} and two from proprioceptors ψ_{m_1}, ψ_{m_2} sensing motor velocity, see Fig. 5.2. We then assume that it only models four hidden states $x = \{l_1, l_2, m_1, m_2\}$, one for each input, see table 5.1.

To specify the agent's generative density necessary in equation (5.2), $p(\psi, x) = p(\psi|x)p(x)$, we then define a likelihood, $p(\psi|x)$ and a prior $p(x)$ in terms of the

[1]The code is available at https://github.com/mbaltieri/braitenberg_vehicles.

TABLE 5.1: **Variables used in the definition of the generative model.**

Variable	Description
ψ	Set of sensory inputs $\{\psi_{l_1}, \psi_{l_2}, \psi_{m_1}, \psi_{m_2}\}$
ψ_{l_1}, ψ_{l_2}	Luminance readings from sensors 1 and 2 (exteroceptors)
ψ_{m_1}, ψ_{m_2}	Velocity readings from motors 1 and 2 (proprioceptors)
x	Set of estimated hidden states $\{x_{l_1}, x_{l_2}, x_{m_1}, x_{m_2}\}$
x_{l_1}, x_{l_2}	Hidden states of exteroceptive sensory readings in the generative model
x_{m_1}, x_{m_2}	Hidden states of proprioceptive sensory readings in the generative model
$\mu_{x_{l_1}}, \mu_{x_{l_2}}$	Means/expected hidden states of exteroceptive sensory readings in the generative model
$\mu_{x_{m_1}}, \mu_{x_{m_2}}$	Means/expected hidden states of proprioceptive sensory readings in the generative model
z, w	Gaussian variables representing uncertainty of the agent on sensory input and dynamics, respectively

agent's estimates of hidden states x. In order to do so we first prescribe a model of how exteroceptive sensations (i.e., light intensity) are generated according to the agent:

$$\psi_{l_1} = x_{l_1} + z_{l_1}, \quad \psi_{l_2} = x_{l_2} + z_{l_2}, \tag{5.5}$$

and similarly for the proprioceptors, representing readings of the velocity of each motor:

$$\psi_{m_1} = x_{m_1} + z_{m_1}, \quad \psi_{m_2} = x_{m_2} + z_{m_2} \tag{5.6}$$

where we effectively assume that sensory readings are linearly related to their hidden states, with some additive zero-mean Gaussian noise $z = \{z_{l_1}, z_{l_2}, z_{m_1}, z_{m_2}\}$ with variance $\sigma_z^2 = \{\sigma_{z_{l_1}}^2, \sigma_{z_{l_2}}^2, \sigma_{z_{m_1}}^2, \sigma_{z_{m_2}}^2\}$ representing measurement uncertainty. The agent's priors on hidden states are then specified in terms of $p(x_m, x_l) = p(x_m|x_l)p(x_l)$, with variables x_{m_1}, x_{m_2} only depending on x_{l_1}, x_{l_2}. This relation implements a very simple and naive assumption for the agent: predictions on proprioceptive input x_{m_1}, x_{m_2} depend upon predictions on luminance (exteroceptive sensations) x_{l_1}, x_{l_2}, but not the other way around. This is a simplifying assumption that, as we will see, enforces a (nearly) reactive motor behaviour triggered by an (almost) input-output response, i.e., proprioceptive input instantly guided by exteroceptive readings of the sensors on the opposite side. We then write a model of the priors as:

$$x_{m_1} = x_{l_2} + w_{m_1}, \quad x_{m_2} = x_{l_1} + w_{m_2} \tag{5.7}$$

where $w = \{w_{m_1}, w_{m_2}\}$ are zero-mean Gaussian noise terms with variance $\sigma_w^2 = \{\sigma_{w_{m_1}}^2, \sigma_{w_{m_2}}^2\}$, describing the uncertainty on the dynamics of the system. Effectively, we specify relations between sensors and motors, describing the underlying dynamics in terms of a contralateral relationship between estimates of exteroceptive (light) sensors x_l and proprioceptive (motor velocities) x_m. These priors will make our agent functionally consistent with Braitenberg vehicle 2b, the "aggressor" (Braitenberg, 1986). We then assume non-informative Gaussian priors on x_l, $p(x_l)$, i.e., they have high variance (low precision), making them nearly uniform and thus removing them from our model for simplicity.

Under the assumption that random variables z are Gaussian with zero mean as in equation (3.31), we can define the likelihood functions

$$\begin{aligned} p(\psi_l|x_l) &= \frac{1}{\sqrt{2\pi\sigma_{z_l}^2}} \exp\left(\frac{-(\psi_l - x_l)^2}{(2\sigma_{z_l}^2)}\right) \\ p(\psi_m|x_m) &= \frac{1}{\sqrt{2\pi\sigma_{z_m}^2}} \exp\left(\frac{-(\psi_m - x_m)^2}{(2\sigma_{z_m}^2)}\right) \end{aligned} \tag{5.8}$$

where $l = \{l_1, l_2\}$ and $m = \{m_1, m_2\}$. Similarly, with Gaussian noise w the priors become

$$\begin{aligned} p(x_{m_1}|x_{l_2}) &= \frac{1}{\sqrt{2\pi\sigma_{w_{m_1}}^2}} \exp\left(\frac{-(x_{m_1} - x_{l_2})^2}{(2\sigma_{w_{m_1}}^2)}\right) \\ p(x_{m_2}|x_{l_1}) &= \frac{1}{\sqrt{2\pi\sigma_{w_{m_2}}^2}} \exp\left(\frac{-(x_{m_2} - x_{l_1})^2}{(2\sigma_{w_{m_2}}^2)}\right) \end{aligned} \tag{5.9}$$

As a result of the Gaussian assumption for both z and w and considering the Laplace assumption on the variational energy introduced in Chapter 3, the free energy is evaluated at the mode of the generative density (see equation (5.2) and Chapter 3) and reduces to (without constants):

$$\begin{aligned} F \approx \frac{1}{2}\Big(&\pi_{z_{l_1}}(\psi_{l_1} - \mu_{l_1})^2 + \pi_{z_{l_2}}(\psi_{l_2} - \mu_{l_2})^2 + \pi_{z_{m_1}}(\psi_{m_1} - \mu_{m_1})^2 + \pi_{z_{m_2}}(\psi_{m_2} - \mu_{m_2})^2 \\ &+ \pi_{w_{m_1}}(\mu_{m_1} - \mu_{l_2})^2 + \pi_{w_{m_2}}(\mu_{m_2} - \mu_{x_1})^2 + \ln(\pi_{z_{l_1}}\pi_{z_{l_2}}\pi_{z_{m_1}}\pi_{z_{m_2}}\pi_{w_{m_1}}\pi_{w_{m_2}})\Big) \end{aligned} \tag{5.10}$$

where $\pi_{z_{l_1}}, \pi_{z_{l_2}}, \pi_{z_{m_1}}, \pi_{z_{m_2}}$ and $\pi_{w_{m_1}}, \pi_{w_{m_2}}$ are the inverse variances of noise terms z and w respectively, also called precisions (see Chapter 3). Precision parameters weight predictions errors based on the agent's confidence on a certain expectation. High precisions imply low variances/uncertainty and thus high confidence, and vice versa. These parameters allow for different emphases on the minimisation of free energy, for example, an agent could focus more on predictions based on its priors and weighted by π_w or rely more on sensations from the environment when π_z are large (see Chapter 4). Some of the implications of this weighting mechanism will be developed in more detail for this model with our simulations. We also explicitly replaced hidden states x with their means μ_x, to highlight the fact that under the Laplace approximation, the conditional variances ς^2 can be obtained analytically as explained in Chapter 3. With the expression for the free energy we can now derive explicit equations for our model that will implement perception (equation (5.3)):

$$
\begin{aligned}
\dot{\mu}_{l_1} &= -k\Big(\pi_{z_{l_1}}(\mu_{l_1} - \psi_{l_1}) + \pi_{w_{m_2}}(\mu_{l_1} - \mu_{m_2})\Big) \\
\dot{\mu}_{l_2} &= -k\Big(\pi_{z_{l_2}}(\mu_{l_2} - \psi_{l_2}) + \pi_{w_{m_1}}(\mu_{l_2} - \mu_{m_1})\Big) \\
\dot{\mu}_{m_1} &= -k\Big(\pi_{z_{m_1}}(\mu_{m_1} - \psi_{m_1}) + \pi_{w_{m_1}}(\mu_{m_1} - \mu_{l_2})\Big) \\
\dot{\mu}_{m_2} &= -k\Big(\pi_{z_{m_1}}(\mu_{m_2} - \psi_{m_2}) + \pi_{w_{m_2}}(\mu_{m_2} - \mu_{l_1})\Big)
\end{aligned}
\tag{5.11}
$$

and action (equation (5.4)):

$$
\begin{aligned}
\dot{a}_1 =& -k\Big(\pi_{z_{m_1}}(\psi_{m_1} - \mu_{m_1})\frac{\partial \psi_{m_1}}{\partial a_1} + \pi_{z_{m_2}}(\psi_{m_2} - \mu_{m_2})\frac{\partial \psi_{m_2}}{\partial a_1}\Big) \\
\dot{a}_2 =& -k\Big(\pi_{z_{m_1}}(\psi_{m_1} - \mu_{m_1})\frac{\partial \psi_{m_1}}{\partial a_2} + \pi_{z_{m_2}}(\psi_{m_2} - \mu_{m_2})\frac{\partial \psi_{m_2}}{\partial a_2}\Big)
\end{aligned}
\tag{5.12}
$$

where k is the learning rate used for the gradient descent in Baltieri and Buckley (2017) and here simply equal to 1.

To implement action according to equation (5.4) we must first introduce the agent's model of the relationship between its actions and its observations. Here this relationship will be defined in a way consistent with the idea of sensorimotor contingencies (Seth, 2014b), i.e., simple mappings between actions and percepts. We assume that actions a_1, a_2 can influence proprioceptive sensations ψ_{m_1}, ψ_{m_2} but, for simplicity, not exteroceptive ones ψ_{l_1}, ψ_{l_2}. This is clearly an unrealistic assumption since actions do, in practice, change all types of sensory inputs. More in general however, this speaks to our broader claim on the importance of simple generative models coupled in the "right way" to their environment over their objective accuracy. Although not the most accurate, these models may constitute a good set of heuristics for the achievement of certain goals in, possibly, simplest way (i.e., Braitenberg-like vehicles). As explained in Friston et al. (2010a), active inference dispenses with the traditional notion of forward/inverse models for motor control in favour of a more general generative (forward) model inverted through Bayesian inference. In this framework, the “inverse” model is thought to be implicitly encoded in predictions about proprioceptive consequences that are implemented through simple reflex arcs embodied in an agent. In our agent, the reflex arcs correspond to the partial the partial derivatives $\partial\psi_{m_{1,2}}/\partial a_{1,2}$ and will be encoded as:

$$\begin{bmatrix} \frac{\partial\psi_{m_1}}{\partial a_1} & \frac{\partial\psi_{m_2}}{\partial a_1} \\ \frac{\partial\psi_{m_1}}{\partial a_2} & \frac{\partial\psi_{m_2}}{\partial a_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.13}$$

Motor reflex arcs are then implemented in the world with the assumption that each motor takes as an input one of the actions produced through the minimisation of free energy

$$v_1 = a_1, \quad v_2 = a_2. \tag{5.14}$$

where v_1, v_2 are the real linear velocities of the vehicle's wheels. One may ask why we are not providing actions as forces (and thus accelerations) for the agent and this is simply due to a simplification we made to avoid the formulation of dynamical models in state-space models in generalised coordinates of motion, used in the following chapters. In general, if the only readings included (wheel) velocities, one would need extra information regarding the mechanisms that transform accelerations in velocities for the vehicle and normally implicitly assumed to be available in implementations of standard controllers.

5.3 Simulations

5.3.1 Phototaxis

Each column in Fig. 5.3 shows the simulation of 10 different vehicles with random initial conditions (position and orientation of the agent). We also added Gaussian noise to each precision parameter (mean varying up to $\pm 20\%$ of the nominal value of each log-precision, variance 1) to confirm the robustness of the solution.

The behaviour of agents operating under the FEP can be shown to depend on the value of precision parameters. In our first simulation, we adjust the precisions to implement phototaxis. Priors in the generative model implement linear contralateral relations between μ_l and μ_m (see equation (5.7)). In this setup, our agent needs to accurately infer μ_l from its readings on luminance ψ_l (Fig. 5.3c top). Expectations μ_l are then mapped to expectations on proprioceptive input μ_m (Fig. 5.3c bottom), see priors in equation (5.7) and projected to the motors via fast (instantaneous in our simulations) reflex arcs, i.e., $\psi_m = \mu_m$. To do so, precisions on proprioceptive sensations (π_{z_m}) need to be much lower relatively to those on the priors (π_{w_m}), in turn smaller than the precisions on exteroceptors (π_{z_l}).

This enables the sensorimotor flow we just implicitly described: 1) inference of exteroceptive sensations, 2) coupling between expectations about light levels and motor velocities and 3) actuation of the motors based on the agent's expectations about proprioceptive input (dictated by the light intensity). In table 5.2, we show

TABLE 5.2: **Simplified equations for phototaxis.** Approximations of equation (5.11) and equation (5.12) with assumptions on precisions enacting phototactic behaviour.

Approximate dynamics - phototaxis	
Left sensor, right motor	Right sensor, left motor
1) $\dot{\mu}_{l_1} \approx -\pi_{z_{l_1}}(\mu_{l_1} - \psi_{l_1})$	1) $\dot{\mu}_{l_2} \approx -\pi_{z_{l_2}}(\mu_{l_2} - \psi_{l_2})$
2) $\dot{\mu}_{m_2} \approx -\pi_{w_{m_2}}(\mu_{m_2} - \mu_{l_1})$	2) $\dot{\mu}_{m_1} \approx -\pi_{w_{m_1}}(\mu_{m_1} - \mu_{l_2})$
3) $\dot{a}_2 \approx -\pi_{z_{m_2}}(\psi_{m_2} - \mu_{m_2})$	3) $\dot{a}_1 \approx -\pi_{z_{m_1}}(\psi_{m_1} - \mu_{m_1})$

how the formulation of perception and action set out in equation (5.11) and equation (5.12), respectively, is simplified by these assumptions for the left-sensor/right-motor relation, $\{\mu_{l_1}, \mu_{m_2}\}$ and the right-sensor/left-motor one $\{\mu_{l_2}, \mu_{m_1}\}$. When these assumptions are met, the vehicle performs phototaxis akin to Braitenberg vehicle 2b (Braitenberg, 1986).

Fig. 5.3a shows the “aggressor-like" behaviour of our agent, which speeds up close to the light and slows down away from it. At the beginning of our simulation we initialise all estimates x to zero. Fig. 5.3e shows how, after a brief transient due to these initial conditions, the free energy rapidly approaches a minimum value exhibiting fluctuations only driven by noise on exteroceptive input. This minimum is reached in less than 0.1 seconds, while it takes the agent nearly 2 seconds to reach the light, see Fig. 5.3c top. This is because the generative model we define does not encode explicit priors on light levels and thus does not specify a target luminance. Instead, the agent minimises free energy by simply satisfying a sensorimotor mapping between expectations about light levels and motor velocities. Phototaxis emerges as a consequence of the coupling between expectations μ_l and μ_m, not because of an explicitly encoded goal.

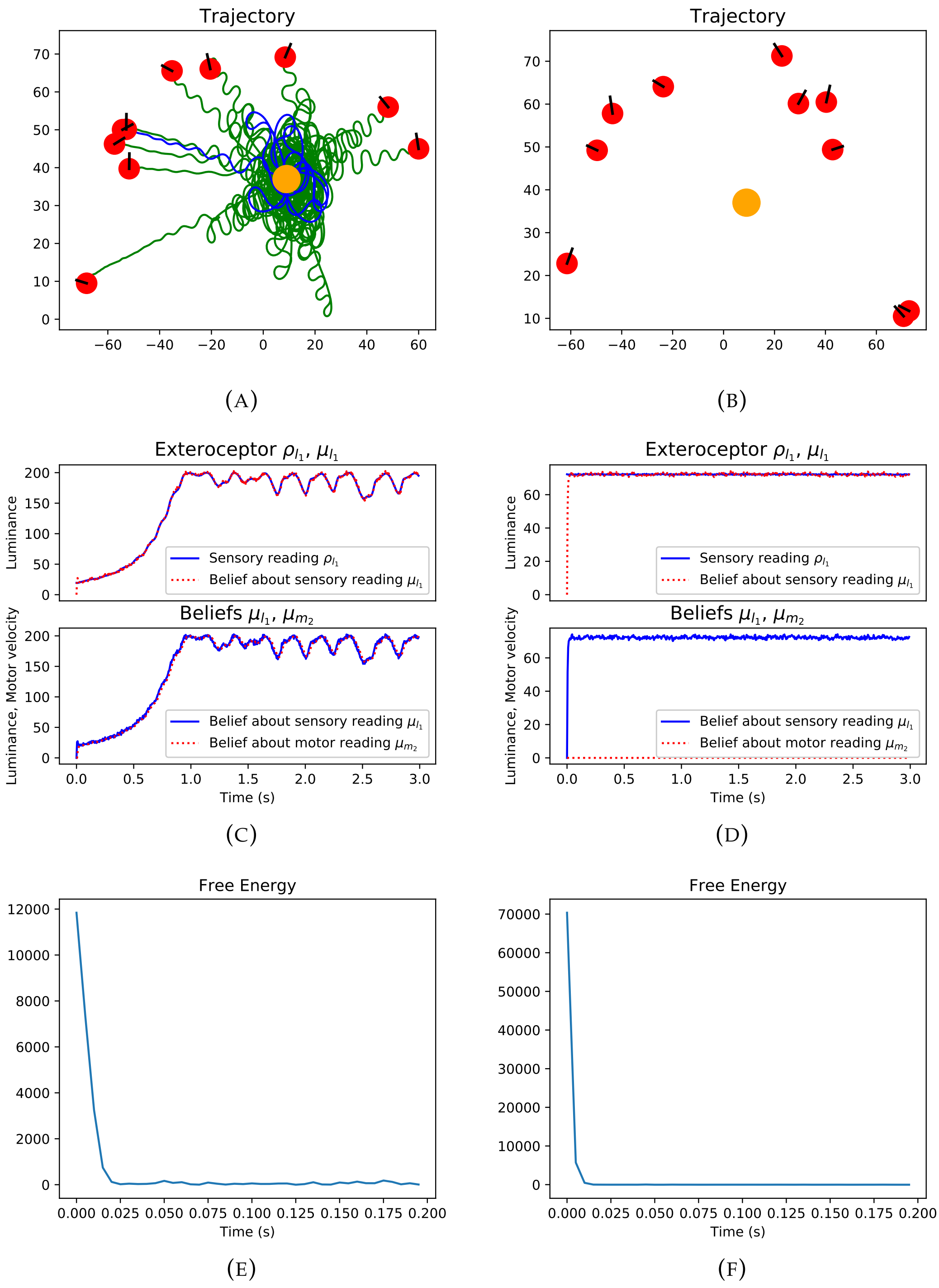


FIGURE 5.3: **Phototaxis and akinesia under active inference.** First column: Phototaxis under the active inference. (A) Trajectory of the vehicle over 5 seconds of simulation. The red circles represent the agents at their initial positions (with orientations parallel to the vertical axis shown using black lines), the yellow circle is the light source. (C - top) True (blue line) and inferred (red line) sensory input ψ_{l_1} and μ_{l_1}. (C - bottom) Coupling of expectation on left sensor μ_{l_1} with expectation on right motor velocity μ_{m_2}. (E) Free energy over the first 0.2 seconds. Second column: Example of "pathological" behaviour under the FEP/Active Inference (akinesia) obtained with higher sensory/proprioceptive precisions. Same layout described above with figures (B), (D), (F).

5.3.2 Pathological behaviour

The sensitivity of behaviour to precisions has been used extensively to develop hypotheses and computational models for phenomena including for instance psychosis and schizophrenia (Adams et al., 2013; Friston et al., 2016b), sensory attenuation (Brown et al., 2013) and attention (Feldman and Friston, 2010). Here we present an interpretation of the role of precisions for our agent inspired by these accounts.

If the balance of precisions on proprioceptive inputs π_{z_m} and on priors π_{w_m} is altered, with an increase of the former and a decrease of the latter, the behaviour of our vehicle is severely affected by the change. The sensorimotor flow necessary for phototaxis is disrupted and while the agent still infers light levels through μ_l (Fig. 5.3d top), these expectations are not mapped to μ_m (Fig. 5.3d bottom) since the agent's precisions about their relation π_{w_m} are dominated by precisions on proprioceptive input π_{z_m} (equation 2 in table 5.3).

This behaviour is consistent with interpretations, under the FEP, of motor control disorders where movements are limited (hypokinesia) or entirely absent (akinesia) (Brown et al., 2013). In active inference terms, decreasing proprioceptive prediction errors precisions, π_{z_m} in our case, is thought to be a necessary condition to actuate motor commands (Brown et al., 2013). Failing to reduce them generates atypical behaviour corresponding to an agent completely unable to move as in the case we discussed when precisions on proprioceptive input π_{z_m} are larger than precisions π_{w_m} (Fig. 5.3b), or having limited motor capabilities when π_{z_m} and π_{w_m} are closer in magnitude (not shown here). More importantly, processes of inference on new incoming sensations are preserved, but this information is not used to trigger appropriate motor commands (see the passive tracker in Chapter 4).

TABLE 5.3: **Simplified equations for pathological behaviour.** Approximations of equation (5.11) and equation (5.12) with assumptions on precisions preventing phototactic behaviour.

Approximate dynamics - pathological behaviour	
Left sensor, right motor	Right sensor, left motor
1) $\dot{\mu}_{l_1} \approx -\pi_{z_{l_1}}(\mu_{l_1} - \psi_{l_1})$	1) $\dot{\mu}_{l_2} \approx -\pi_{z_{l_2}}(\mu_{l_2} - \psi_{l_2})$
2) $\dot{\mu}_{m_2} \approx -\pi_{z_{m_2}}(\mu_{m_2} - \psi_{m_2})$	2) $\dot{\mu}_{m_1} \approx -\pi_{z_{m_1}}(\mu_{m_1} - \psi_{m_1})$
3) $\dot{a}_2 \approx -\pi_{z_{m_2}}(\psi_{m_2} - \mu_{m_2})$	3) $\dot{a}_1 \approx -\pi_{z_{m_1}}(\psi_{m_1} - \mu_{m_1})$

5.3.3 Other vehicles

Previously, we showed that different precisions within the same generative model can qualitatively affect the behaviour of an agent, going from performing phototaxis to a catatonic state by just regulating some of these weights. In this section we explore how different priors, encoding in our case the relationships between exteroceptive and proprioceptive inputs (see equation (5.7)) can also be used to generate new emergent behaviour. As we saw in section 5.2, during the definition of our generative model, vehicles in Fig. 5.3a described trajectories consistent with Braitenberg vehicle 2b. Simple changes to the priors in equation (5.7) can easily reproduce behaviours analogous to other vehicles, for instance 2a (the “coward"), 3a (the “lover") and 3b (the “explorer").

In more detail, models of the priors were updated to:

- for the coward vehicle, rapidly running away from the light to then slow down once distance is gained,

$$\mu_{m_1} = \mu_{l_1} + w_{m_1}, \quad \mu_{m_2} = \mu_{l_2} + w_{m_2} \tag{5.15}$$

- for the lover vehicle, quickly approaching the light while decreasing its speed close to the source,

$$\mu_{m_1} = l_{max} - \mu_{l_1} + w_{m_1}, \quad \mu_{m_2} = l_{max} - \mu_{l_2} + w_{m_2} \tag{5.16}$$

- for the explorer vehicle, repelled by the light, moving slowly in its vicinity and speeding up away from it,

$$\mu_{m_1} = l_{max} - \mu_{l_2} + w_{m_1}, \quad \mu_{m_2} = l_{max} - \mu_{l_1} + w_{m_2} \tag{5.17}$$

with l_{max} representing the highest light intensity in the agents' environment.

5.4 Discussion

In this chapter, we presented an implementation of phototaxis within an emerging framework in computational and cognitive neuroscience, the Free Energy Principle (FEP). According to the FEP, processes like perception, learning and action can be defined in biological systems as the minimisation of a quantity defined as (variational) free energy (Friston, Kilner, and Harrison, 2006; Friston, 2010b; Friston, 2012). Most of the simulations proposed so far have focused on “perception-centric" interpretations of the FEP, where generative models play the role of accurate descriptions of the dynamics of the world they represent. These models can capture the intrinsic properties of an agent's sensations and reconstruct the hidden properties of these sensations to a great degree of detail. On this view, a model is assessed on how accurately it can predict incoming sensations, behaviour only emerges as a consequence of reliable information encoded by the agent.

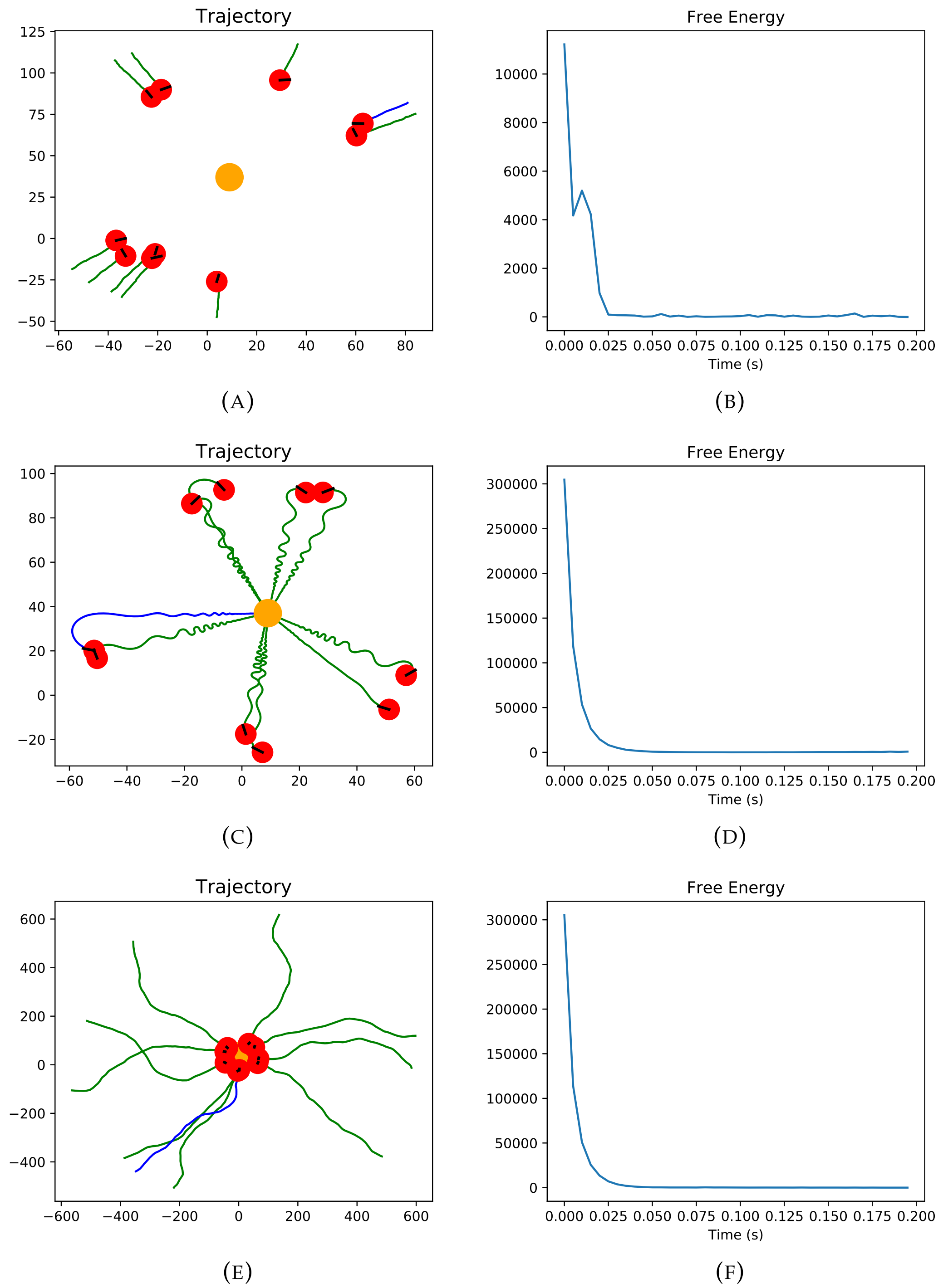


FIGURE 5.4: **Different braitenberg vehicles are obtained by updating the priors of the agent.** (A) shows the "coward"-like behaviour of Braitenberg vehicle 2a. (C) is akin to vehicle 3a, the "lover". (E) behaves as vehicle 3b, the "explorer". On the right, (B), (D) and (F) show the respective processes of minimisation of variational free energy. Ten simulations were performed for each set up, with random initial conditions (position and orientation of the agent). We also added Gaussian noise to each precision parameter (mean varying up to $\pm 20\%$ of the value of each precision parameter, variance 1) to confirm the robustness of the solution.

Our implementation, on the other hand, represents an example of "action-oriented" interpretations of the FEP framework (Seth, 2014b; Clark, 2015a; Bruineberg, Kiverstein, and Rietveld, 2018; Allen and Friston, 2018), with a focus on minimal generative models. In this case, models are evaluated based on their ability to allow an agent to perform a task or achieve a certain goal. The information recapitulated in these models is only relevant to an agent engaging in ecologically relevant behaviour: inferring and encoding more properties of the world dynamics is not necessarily an advantage (Ashby, 1958). This idea is described within the formalism of the FEP as the complexity-accuracy trade off. It is possible, in fact, to rewrite equation (3.6) in information theoretical terms as a combination of measures of complexity and accuracy (Friston, Kilner, and Harrison, 2006; Friston, 2010b), see equation (3.10). According to this alternative reading, minimising variational free energy is equivalent to maximising a measure of the predictive power (i.e., accuracy) of a generative model while minimising its complexity (i.e., the number of degrees of freedom). An agent, according to the FEP, is then mandated to encode only relevant information to avoid unnecessary complex models that don't improve its performances.

Our agent's generative model contains a set of variables, $\mu_x = \{\mu_{l_1}, \mu_{l_2}, \mu_{m_1}, \mu_{m_2}\}$ acting as expectations on the hidden states x of sensations $\psi = \{\psi_{l_1}, \psi_{l_2}, \psi_{m_1}, \psi_{m_2}\}$. In our implementation, these variables are not strongly related to how sensory data ψ's are actually generated, i.e., they contain no information about, for example, the details of the agent's body or its environment. It could be argued that this agent does not even possess a *generative* model since, effectively, it cannot generate predictions in line with sensory data if not tightly coupled to its environment (cf. classical ideas such as the wake phase in the wake-sleep algorithm (Dayan et al., 1995)). Our interpretation, however, aligns with (Clark, 2015a; Bruineberg, Kiverstein, and Rietveld, 2018) in saying that a generative model in the context of embodied agents should be assessed based on its ability to prescribe an agent to perform a task (action-oriented), rather than on how well it can reconstruct and accurately predict sensory input (perception-centric).

In our simulations, we first provided an account of phototaxis functionally consistent with vehicle 2b (Braitenberg, 1986), and how it could be implemented under the FEP. To allow for such behaviour, our agent was endowed with expectations on exteroceptive readings (i.e., light intensity) mapped to proprioceptive ones (i.e., motor velocity) using simple linear priors on contralateral relations. In addition, this agent needs to implement high and low precisions on exteroceptive and proprioceptive inputs respectively, with precisions on their interaction placed somewhere in between. This agent performs phototaxis through active inference without explicitly encoding information about a target end-point in the generative model, i.e., it does not specify a light intensity to achieve. The priors of this agent encode a target "state of affairs" rather than a final goal. This agent thus minimises its free energy by complying with this state of affairs, not by achieving an explicit goal state, similarly to the idea of homeokinesis (Der and Martius, 2012). Phototaxis is just a consequence of sensorimotor contingencies represented by how priors relate light levels to motor velocities.

We then explored behaviours defined as “pathological", inspired by work on the FEP for motor disorders, psychosis and schizophrenia among others (Adams et al., 2013; Friston et al., 2016b). Some pathologies, it has been suggested, can be recapitulated under the FEP and active inference using different weighting parameters in a generative model (i.e., precisions). Our simulations investigated a case where a combination of high precisions on proprioceptive prediction errors and low precisions on the prior coupling between extero- and proprioceptive expectations resulted in a reduced (or complete lack of) ability to move. A similar idea was presented by Brown et al. (2013) where decreasing the confidence (precision) of sensations about self-generated movements is thought to be a necessary condition for the initiation of action. We finally explored how simple changes in priors can produce new emergent behaviours, for example an alternative version of phototaxis and two types of photophobia, in line with vehicles 2a and 3a-b (Braitenberg, 1986), .

In the future, we will investigate the possible functional benefits of this architecture based on the FEP. For instance, preliminary results comparing standard Braitenberg vehicles and our implementation already suggest a higher robustness of the active inference agent in very noisy environments because of low-pass filters implemented by equation (5.11) and equation (5.12). A second direction will address generative models with a less minimal set of assumptions, including for example priors on different exteroceptive inputs (e.g., a target position to reach?) or interoceptive ones (e.g., the maintenance of temperature or other variables related to luminance?). These priors could model the exploration of behaviours that can only be performed with the implementation of such mechanisms (e.g., dynamic trajectories such as keeping a certain distance from the light source while maintaining a certain temperature?). This will allow us to investigate the implications for the perception-centric vs. action-oriented debate on the interpretation of the FEP (Clark, 2015a; Bruineberg, Kiverstein, and Rietveld, 2018; Allen and Friston, 2018) for more complex adaptive behaviours. It would also be interesting to investigate the “wiggling” behaviour displayed by all of our agents, see Fig. 5.3 and Fig. 5.4. This is likely to be appear because of the small delays introduced between the estimation of light levels and the implementation of motor actions in our numerical simulations (smaller time steps were already reducing the jiggling, although not cancelling it out completely). In active inference we could dispose of this abnormal behaviour with the implementation of delayed differential equations using generalised coordinates of motion in our model, discounting for these small delays by including them into the agent's generative model, see for instance Perrinet, Adams, and Friston (2014).

5.5 Conclusions

The implementation of phototaxis we presented here is admittedly rather complicated. However, this chapter constitutes a simple but complete proof of concept of action-oriented approaches to the FEP. By creating a generative model encoding virtually no objective knowledge of the environment, we showed that the common intuition prescribing the necessity of agents as mirrors of their world in the FEP and other Bayesian accounts of cognition is misguided. In this chapter's set up, behaviour can emerge as a simple consequence of strong priors enforcing a tight coupling between an agent and its milieu. The minimisation of variational free energy describes then only the realisation of such coupling, while the dynamics of an agent-environment system altogether generate (apparently) purposeful behaviour as in classic implementations of 4E ideas, e.g., Braitenberg (1986). The action-oriented model we implemented has also begun to address some of the concerns about the FEP exposed in Clark (2015a) and Bruineberg, Kiverstein, and Rietveld (2018), essentially worrying that interpretations of this framework have been so far too dependent on a perception-centric account of behaviour. Our example shows that this need not be the case. We would argue that, more in general, the FEP is neutral on the implementation details of agentive behaviour and is thus compatible with many different cognitive frameworks, including 4E.

Chapter 6

An active inference models of robust regulation via integral control

The relationship between information/probability theory and control theory has long been recognised, with the first intuitions emerging from work by Ashby (1958), Shannon (1959), and Kalman (1960c). As discussed in Chapter 2, a unifying view of these two theoretical frameworks is nowadays proposed for instance in stochastic optimal control (Stengel, 1994; Todorov, 2008) and extended in active inference (Friston, 2011), with connections to ideas of sensorimotor loops in biological systems (Friston et al., 2010a; Friston et al., 2015). These connections emphasise homeostasis, regulation and concepts such as set-point control and negative feedback for the study of different aspects of living systems, with roots in the cybernetics movement (Ashby, 1957; Wiener, 1961). It remains, however, unclear how the active inference formulation directly relates to more traditional concepts of classical control theory. PID control, a popular control strategy working with little prior knowledge of the process to regulate, is commonly applied in engineering (Åström, 1995; Ang, Chong, and Li, 2005; Åström and Hägglund, 2006) and more recently used in biology and neuroscience modelling (Yi et al., 2000; Yang and Iglesias, 2006; Ang et al., 2010; Ritz et al., 2018; Chevalier et al., 2018). In this chapter, we develop an information theoretic interpretation of PID control, showing how it can be derived in a more general Bayesian (active) inference framework. We will show, in particular, that approximate models of the world are often enough for regulation, and in particular that simple linear generative models that only approximate the true dynamics of the environment implement PID control as a process of inference. Using this formulation we also propose a new method for the optimisation of the gains of PID controllers based on the same principle of variational free energy minimisation, and implemented as a second order optimisation process. Finally, we will show that our implementation of PID controllers as approximate Bayesian inference lends itself to a general framework for the formalisation of different (conflicting) criteria in the design of a controller, the so-called performance-robustness trade-off (Åström and Hägglund, 2001; Åström and Hägglund, 2006), as a cohesive set of constraints implemented in a free energy functional. In active inference, these criteria will be mapped to precisions, or inverse variances, of observations and dynamics of a state-space model with a straightforward interpretation in terms of uncertainty on different variables of a system.

This chapter extends results from Chapter 5 and is based on Baltieri and Buckley (2018a) and Baltieri and Buckley (2019c). Here we emphasise once more the role of potential differences between the generative model of an agent and the generative process of its environment. The generative model in particular is a simple linear approximation of the world dynamics with priors representing the agent's target state (or more in general target trajectory). Regulation equivalent to PID control is achieved by agents forcing the world to match their desires expressed in terms of linear dynamics. The chapter includes a mathematical derivation PID-like control in active inference, showing then the equivalence between the parameters or gains of a PID controller and precisions of the FEP formulation. I then propose an analytical mechanism for the optimisation of these gains based on the optimisation of precision hyperparameters introduced in Chapter 3. An agent-based model of cruise control with more realistic and complex dynamics (cf. Chapter 3) is then used to simulate our active inference agent and test its robustness in presence of disturbances, changes to target states and varying observation noise. We finally test the optimisation of (hyper)parameters, showing the improvements achieved by our PID-like agent.

6.1 PID control

Proportional-Integral-Derivative (PID) control is one of the most popular types of controllers used in industrial applications, with more than 90% of total controllers implementing PID or PI (no derivative) regulation (Åström and Hägglund, 2004; Åström and Hägglund, 2006). It is one of the simplest set-point regulators, whereby a desired state (i.e., set-point, reference, target) represents the final goal of the regulation process, e.g., to maintain a room temperature of 23° C. PID controllers are based on closed-loop strategies with a negative feedback mechanism that tracks the real state of the environment. In the most traditional implementation of negative feedback methods, the difference between the measured state of the variable to regulate (e.g., the real temperature in a room) and the target value (e.g., 23° C) produces a prediction error whose minimisation drives the controller's output, e.g., if the temperature is too high, it is decreased and if too low, it is raised. In mathematical terms:

$$e(t) = y_r - y(t) \tag{6.1}$$

where $e(t)$ is the error, y_r is the *reference* or set-point (e.g., desired temperature) and $y(t)$ is the observed variable (e.g., the actual room temperature).

This mechanism is, however, unstable in very common conditions, in particular when a steady-state offset is added (e.g., a sudden and unpredictable change in external conditions affecting the room temperature which are not under our control), or when fluctuations need to be suppressed (e.g., too many oscillations while regulating the temperature may be undesirable). PID controllers elegantly deal with both of these problems by augmenting the standard negative feedback architecture, here called *proportional* or *P term*, with an *integral* or *I* and a *derivative* or *D term*, see Fig. 6.1. The integral term accumulates the prediction error over time in order to cancel out errors due to unaccounted steady-state input, while minimising the derivative of the prediction error leads to a decrease in the amplitude of fluctuations of the controlled signal. The general form of the control signal $u(t)$ generated by a PID controller is usually described by:

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau) d\tau + k_d \frac{de(t)}{dt} \tag{6.2}$$

where $e(t)$ is, again, the prediction error and k_p, k_i, k_d are the so called proportional, integral and derivative gains respectively, a set of parameters used to tune the relative strength of the P, I and D terms of the controller.

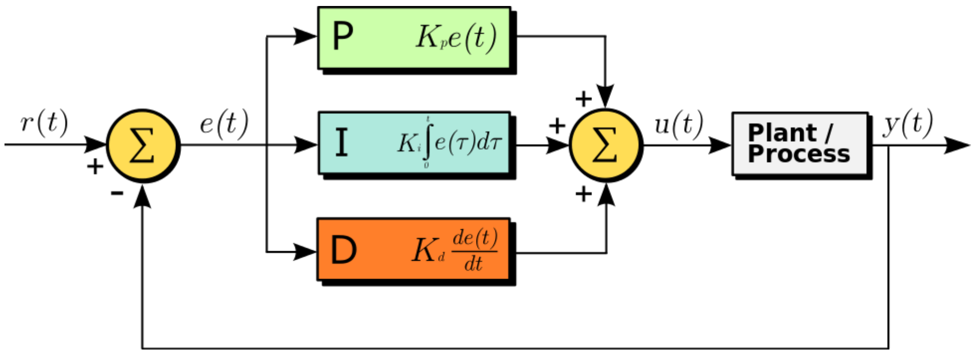


FIGURE 6.1: **A PID controller.** The prediction error $e(t)$ is given by the difference between a reference signal $r(t)$, y_r in our formulation, and the output $y(t)$ of a process. The different terms, one proportional to the error (P term), one integrating the error over time (I term) and one differentiating it (D term), drive the control signal $u(t)$. Image taken from (Arturo Urquizo, 2011).

The popularity of PID controllers is largely due to their robustness in the presence of uncertainty, i.e., step disturbances and more in general measurement noise, given by the filtering properties of the I term. One of the major challenges on the other hand, lies with the tuning of parameters k_p, k_i, k_d, that have to be adapted to deal with different (often conflicting) constraints on the regulation process (Åström, 1995; Åström and Hägglund, 2001).

6.1.1 The performance-robustness trade-off

The presence of conflicting criteria for the design of PID controller is a well known issue in the control theory literature, often referred to as the performance-robustness trade-off (Rivera, Morari, and Skogestad, 1986; Åström, Panagopoulos, and Hägglund, 1998; Åström and Hägglund, 2001; Åström and Hägglund, 2006; Garpinger, Hägglund, and Åström, 2014). A controller needs to optimise pre-specified performance criteria while, at the same time, preserving some level of robustness in face of uncertainty and unexpected conditions during the regulation process. In recent attempts to formalise and standardise these general principles (Åström and Hägglund, 2001; Åström and Hägglund, 2006), the performance of a controller has been proposed to be evaluated through:

- load disturbance response, how a controller reacts to changes in external inputs, e.g., a step input,
- set-point response, how a controller responds to different set-points over time,
- measurement noise response, how noise on the observations impacts the regulation process,

while robustness to be assessed on:

- robustness to model uncertainty, how uncertainty on the plant/environment dynamics affects the controller.

The goal of a general methodology for the design and tuning of PID controllers is to bring together these (and possibly more) criteria into a formal and tractable framework that can be used for a large class of problems. To date, the only proposal attempting to do so is presented in Åström, Panagopoulos, and Hägglund (1998). This methodology is based on the maximisation of the integral gain (equivalent to the minimisation of the integral of the error from the set-point, see Åström (1995)), subject to constraints derived from a frequency domain analysis related to the Nyquist stability criterion applied to the controlled system (Åström, Panagopoulos, and Hägglund, 1998). In this chapter, we propose our formulation also as a general framework for the design and tuning of PID controllers leveraging the straightforward interpretation of the performance-robustness trade-off for PID controllers in terms of uncertainty parameters (i.e., precisions or inverse variances) in the variational free energy.

6.2 PID control as active inference

To implement PID control as a process of active inference, we will first describe an agent's generative model as a generalised linear state space model of second order (i.e., only two higher orders of motion, anything beyond that is zero-mean Gaussian noise):

$$\begin{aligned}
\psi &= x + z & \dot{x} &= x' = -\alpha(x + v) + w \\
\psi' &= x' + z' & \dot{x}' &= x'' = -\alpha(x' + v') + w' \\
\psi'' &= x'' + z'' & \dot{x}'' &= x''' = -\alpha(x'' + v'') + w''
\end{aligned}$$

where $\theta = \alpha$ is a parameter (see Chapter 3). As previously suggested, with a Gaussian assumption on $\tilde{z}, \tilde{w}$, the likelihood is reduced to:

$$P(\tilde{\psi}|\tilde{x}, \tilde{v}; \theta, \gamma) = P(\tilde{\psi}|\tilde{x}; \theta, \gamma) = N(\tilde{\mu}_x, \sigma^2_{\tilde{z}}) \tag{6.3}$$

where we assume no direct dependence of observations $\tilde{\psi}$ on external inputs $\tilde{v}$, while the prior is described by:

$$P(\tilde{x}, \tilde{v}; \theta, \gamma) = P(\tilde{x}|\tilde{v}; \theta, \gamma) P(\tilde{v}; \theta, \gamma) \tag{6.4}$$

with

$$\begin{aligned}
P(\tilde{x}|\tilde{v}; \theta, \gamma) &= N(-\alpha(\tilde{\mu}_x + \tilde{\mu}_v), \sigma^2_{\tilde{w}}) \\
P(\tilde{v}; \theta, \gamma) &= N(\tilde{\eta}_x, \sigma^2_{\tilde{v}})
\end{aligned} \tag{6.5}$$

The Laplace-encoded variational free energy in equation (3.33) then becomes:

$$\begin{aligned}
F \approx \frac{1}{2}\Big[&\pi_z\Big(\psi - \mu_x\Big)^2 + \pi_{z'}\Big(\psi' - \mu'_x\Big)^2 + \pi_{z''}\Big(\psi'' - \mu''_x\Big)^2 + \pi_w\Big(\mu'_x + \alpha(\mu_x - \eta_x)\Big)^2 + \\
&+ \pi_{w'}\Big(\mu''_x + \alpha(\mu'_x - \eta'_x)\Big)^2 + \pi_{w''}\Big(\mu'''_x + \alpha(\mu''_x - \eta''_x)\Big)^2 - \ln\left(\pi_z \pi_w \pi_{z'} \pi_{w'} \pi_{z''} \pi_{w''}\right)\Big]
\end{aligned} \tag{6.6}$$

To simplify our formulation, we assume that precisions $\pi_{\tilde{v}}$ tend to infinity (i.e., no uncertainty on desires), so that $P(\tilde{v}; \theta, \gamma)$ in equation (6.5) becomes a delta function and inputs $\tilde{v}$ reduce to their prior expectations $\tilde{\eta}_x$, i.e., $\tilde{\mu}_v = \tilde{\eta}_x$. With this simplification, prior precisions $\pi_{\tilde{v}}$ and respective predictions errors $(\tilde{\mu}_v - \tilde{\eta}_x)$ are not included in our formulation (see Friston, Trujillo-Barreto, and Daunizeau (2008) and Friston et al. (2010b) for more general treatments). By applying the gradient descent described in Chapter 3 to our free energy functional, we then get the following update equations for perception (estimation):

$$\begin{aligned}
\dot{\mu}_x =& \mu'_x - \Bigg[- \pi_z \Big(\psi - \mu_x \Big) + \pi_w \alpha \Big(\mu'_x + \alpha(\mu_x - \eta_x) \Big) \Bigg] \\
\dot{\mu}'_x =& \mu''_x - \Bigg[- \pi_{z'} \Big(\psi' - \mu'_x \Big) + \pi_{w'} \alpha \Big(\mu''_x + \alpha(\mu'_x - \eta'_x) \Big) + \pi_w \Big(\mu'_x + \alpha(\mu_x - \eta_x) \Big) \Bigg] \\
\dot{\mu}''_x =& \mu'''_x - \Bigg[- \pi_{z''} \Big(\psi'' - \mu''_x \Big) + \pi_{w''} \alpha \Big(\mu'''_x + \alpha(\mu''_x - \eta''_x) \Big) + \pi_{w'} \Big(\mu''_x + \alpha(\mu'_x - \eta'_x) \Big) \Bigg]
\end{aligned} \tag{6.7}$$

and for action (control):

$$\dot{a} = - \Bigg[\pi_z \Big(\psi - \mu_x \Big) \frac{\partial \psi}{\partial a} + \pi_{z'} \Big(\psi' - \mu'_x \Big) \frac{\partial \psi'}{\partial a} + \pi_{z''} \Big(\psi'' - \mu''_x \Big) \frac{\partial \psi''}{\partial a} \Bigg] \tag{6.8}$$

The mapping of these equations to a PID control scheme becomes more clear under a few simplifying assumptions. First, we assume strong priors on the causes of proprioceptive observations. Intuitively, these priors are used to define actions that change the observations to better fit the agent's desires, i.e., the target of the PID controller. This is implemented in the weighting mechanism of prediction errors by precisions in equation (6.6); see also Friston et al. (2010a) and Brown et al. (2013) and Chapter 4 for similar discussions on the role of precisions for behaviour. In our derivation, weighted prediction errors on system dynamics, $\pi_{\tilde{w}}(\tilde{\mu}'_x + \tilde{\mu}_x - \tilde{\eta}_x)$, will be weighted more than weighted errors on observations, $\pi_{\tilde{z}}(\tilde{\psi} - \tilde{\mu}_x)$. To achieve this, we decrease sensory precisions $\pi_{\tilde{z}}$ on proprioceptive observations, effectively biasing the gradient descent procedure towards minimising errors on the prior dynamics (Brown et al., 2013). Secondly, we set the decay parameter α to a large value (theoretically $\alpha \to \infty$, in practice $\alpha = 10^5$ in our simulations), obtaining a set of differential equations including only terms of order α^2 for perception:

$$\begin{aligned}
\dot{\mu}_x \approx& - \pi_w \alpha \Big(\alpha(\mu_x - \eta_x) \Big) \\
\dot{\mu}'_x \approx& - \pi_{w'} \alpha \Big(\alpha(\mu'_x - \eta'_x) \Big) \\
\dot{\mu}''_x \approx& - \pi_{w''} \alpha \Big(\alpha(\mu''_x - \eta''_x) \Big)
\end{aligned} \tag{6.9}$$

This can be interpreted as an agent encoding beliefs in a world that quickly settles to a desired equilibrium state. This assumption effectively decouples orders of generalised motion, with higher orders not affecting the minimisation of lower ones in equation (6.7), since terms from lower orders are modulated by α directly. The remaining terms effectively impose constraints on the generalised motion only close to equilibrium, playing a minor role in the control process away from the target/equilibrium (the more interesting part of regulation). These terms are necessary for the system to settle to a proper steady state when $(\tilde{\mu}_x - \tilde{\eta}_x) \to 0$ and maintain consistency across generalised orders of motion for small fluctuations at steady state, but have virtually no influence at all in conditions far from equilibrium. Following equation (6.9), at steady state, expectations on hidden states $\tilde{\mu}_x$ are mainly driven by priors $\tilde{\eta}_x$:

$$\tilde{\mu}_x = \tilde{\eta}_x \tag{6.10}$$

but are still not met by appropriate changes in observations $\tilde{\psi}$ which effectively implement the regulation around the desired target. To minimise free energy in presence of strong priors, this agent will necessarily have to modify its observations $\tilde{\psi}$ to better match expectations $\tilde{\mu}_x$, which in turn are shaped by priors (i.e., desires) $\tilde{\eta}_x$. Effectively, the agent "imposes" its desires on the world, acting to minimise the prediction errors arising at the proprioceptive sensory layers. In essence, an active inference agent implements set-point regulation by behaving to make its sensations accord with its strong priors/desires. After these assumptions, action can be written as:

$$\dot{a} \approx -\left[\pi_z\Big(\psi - \eta_x\Big)\frac{\partial \psi}{\partial a} + \pi_{z'}\Big(\psi' - \eta_x'\Big)\frac{\partial \psi'}{\partial a} + \pi_{z''}\Big(\psi'' - \eta_x''\Big)\frac{\partial \psi''}{\partial a}\right] \tag{6.11}$$

where we still need to specify partial derivatives $\partial\tilde{\psi}/\partial a$. As discussed in Friston et al. (2010a) and previous chapters, this step highlights the fundamental differences between the FEP and the more traditional forward/inverse models formulation of control problems in biological systems (Kawato, 1999; Wolpert and Ghahramani, 2000). While these derivatives help in the definition of an inverse model (i.e., finding the correct action for a desired output), unlike more traditional approaches, active inference does not involve a mapping from hidden states $\tilde{x}$ to actions a, but is cast in terms of sensory data $\tilde{\psi}$ directly, This is thought to simplify the problem: from a mapping between *unknown* hidden states to actions, to a mapping between *known* observations $\tilde{\psi}$ and actions a. It is claimed that this provides an easier implementation for an inverse model (Friston, 2011), one that is grounded in an extrinsic frame of reference, i.e., the real world ($\tilde{\psi}$), rather than in a intrinsic one in terms of hidden states ($\tilde{x}$) to be inferred first. To achieve PID-like control, we assume that the agent adopts the simplest (i.e., linear) relationship between its actions (controls) and their effects on sensory input across all orders of motion:

$$\frac{\partial \psi}{\partial a} = \frac{\partial \psi'}{\partial a} = \frac{\partial \psi''}{\partial a} = 1 \tag{6.12}$$

This reflects a very simple reflex-arc-like mechanism that is triggered every time a proprioceptive prediction is generated: positive actions (linearly) increase the values of the sensed variables $\tilde{\psi}$, while negative actions decrease them. There is, however, an apparent inconsistency here that we need to dissolve: the proprioceptive input ψ and its higher order states ψ', ψ'' are *all* linearly dependent with respect to actions a as represented in equation (6.12). While an action may not change position, velocity and acceleration of a variable in the same way, a generative model doesn't need to perfectly describe the system to regulate: these derivatives only encode sensorimotor dependencies that allow for, in this case, sub-optimal control. In the same way, PID controllers are, in most cases, effective but only approximate solutions for control (Åström, 1995; Åström and Murray, 2010). This allows us to understand the encoding of an inverse model from the perspective of an agent (i.e., the controller) rather than assuming a perfect, objective mapping from sensations to actions that reflects exactly how actions affect sensory input (Friston et al., 2010a). This also points at possible investigations of generative/inverse models in simpler living systems where accurate internal models are not perhaps needed, and where strategies like PID control are implemented (Yi et al., 2000; Yang and Iglesias, 2006; Ang et al., 2010). By combining equation (6.11) and equation (6.12), action can then be simplified to:

$$\dot{a} \approx \pi_z \big(\eta_x - \psi\big) + \pi_{z'} \big(\eta'_x - \psi'\big) + \pi_{z''} \big(\eta''_x - \psi''\big) \tag{6.13}$$

which is consistent with the "velocity form" or algorithm of a PID controller (Åström, 1995):

$$\dot{u} = k_i \big(y_r - y\big) + k_p \frac{d}{dt} \big(y_r - y\big) + k_d \frac{d^2}{dt^2} \big(y_r - y\big) \tag{6.14}$$

Velocity forms are used in control problems where, for instance, integration is provided by an external mechanism outside the controller (Åström, 1995; Åström and Murray, 2010). Furthermore, velocity algorithms are the most natural form for the implementation of integral control to avoid windup effects of the integral term, emerging when actuators can't regulate an indiscriminate accumulation of steady-state error in the integral term due to physical limitations (Åström, 1995; Svrcek, Mahoney, Young, et al., 2006). This algorithm is usually described using discrete systems to avoid the definition of the derivative of random variables, often assumed to be white noise in the Ito's sense (i.e., Markov processes). In the continuous case, if the variable y is a Markov process, its time derivative is in fact not well defined. For this form to exist in continuous systems, y must be a smooth (stochastic) process. Effectively, this drops the Markov assumption of white noise and implements the same definition of analytic (i.e., differentiable) noise related to the generalised coordinates of motion we described earlier. The presence of extra prediction errors beyond the traditional negative feedback (proportional term) can, in this light, be seen as a natural consequence of considering linear non-Markovian processes with simple reflex mechanisms responding to position, velocity and acceleration in the generalised motion phase space (see equation (6.12)). To ensure that the active inference implementation approximates the velocity form of PID control we still need to clarify the relationship between the generalised coordinates of motion in equation (6.13) and the differential operators $d/dt, d^2/dt^2$ in equation (6.14). As pointed out in previous work, when the variational free energy is minimised, the two of them are equal since the motion of the mode becomes the mode of the motion (Friston, Trujillo-Barreto, and Daunizeau, 2008; Buckley et al., 2017). To simplify our formulation and show PID control more directly, we can consider the case for $\eta'_x = \eta''_x = 0$, defining the more standard set-point control where a desired or set-trajectory collapses to a single set-point in the state-space (cf. homeostasis as homeokinesis collapsed to a single point (Der and Martius, 2012)) and equivalent, in the velocity form, to the case where y_r is a constant and $dy_r/dt = d^2y_r/dt^2 = 0$.

6.3 A model of cruise control

To show an implementation of PID control through active inference we use a standard model of cruise control, i.e., a car trying to maintain a certain velocity over time[1]. While only a toy model, the intuitions and results we derive can easily be transferred to the regulation of proteins in bacterial chemotaxis (Yi et al., 2000) or yeast osmoregulation (Muzzey et al., 2009), and more generally to any homeostatic mechanism (Ashby, 1957), especially when including limited knowledge of external forces (Sontag, 2003). In this setup, a controller receives the speed of the car as an input and adapts the throttle of the vehicle based on a negative feedback mechanism to achieve the desired, or target, cruising velocity. In real-world scenarios, this mechanism needs to be robust in presence of external disturbances, essentially represented by changes in the slope of the road, wind blowing, etc., see Fig. 6.2d. For simplicity, we will use the model based on the formulation in (Åström and Murray, 2010).

The equation of motion of the car is:

$$m\frac{d^2 s}{dt^2} = F - F_d \tag{6.15}$$

where s is the position, F the force generated by the engine and F_d a disturbance force that accounts for a gravitational component F_g, a rolling friction F_r and an aerodynamic drag F_a, such that $F_d = F_g + F_r + F_a$, see again Fig. 6.2d. The forces will be modelled as following:

$$\begin{aligned}
F =& r_g a(t) T_m \left(1 - \beta\left(\frac{\omega}{\omega_m} - 1\right)^2\right)\\
F_g =& mg \sin\lambda\\
F_r =& mg\, C_r \operatorname{sgn} \dot{s}\\
F_a =& \frac{1}{2}\rho\, C_d A \dot{s}^2
\end{aligned} \tag{6.16}$$

with all the constants and variables reported and explained in table 6.1.

In this particular instance, we will simplify PID to PI control since the derivative term is often not used for the cruise control problem (Åström and Murray, 2010), hence implementing only a first order generalised state-space model (see

[1] The code is available at https://github.com/mbaltieri/PIDControlActiveInferenceFEP.

TABLE 6.1: **Cruise control problem, constants and variables.**

	Description	Value
$s(t)$	car position	-
r_g	gear ratio divided by wheel radius	12
$a(t)$	control	-
T_m	maximum torque	$190Nm$
β	motor constant	0.4
ω	engine speed	$\alpha_n v$
ω_m	speed that gives maximum torque	$420rad/s$
m	car mass	$100kg$
g	gravitational acceleration	$9.81m/s^2$
λ	slope of the road	$4°$
C_r	coefficient of rolling friction	0.01
ρ	density of the air	$1.3kg/m^3$
C_d	aerodynamic drag coefficient	0.32
A	frontal area of the car	$2.4m^2$

equation (6.3)). The controller receives noisy readings ψ, ψ' of the true velocity and acceleration of the car, x, x', following the formulation in equation (6.3). The controller is provided with a Gaussian prior in generalised coordinates encoding desired velocity and acceleration with means $\eta_x = 10$ km/h, $\eta'_x = 0$ km/h^2. This prior represents a target trajectory for the agent that, as we aW in equation (6.13), will be equivalent to integral and proportional terms of a PI controller in velocity form. The recognition dynamics are then specified in equation (6.7) and equation (6.8).

In Fig. 6.2 we show the behaviour of a standard simulation of active inference implementing PI-like control for the controller of the speed of a car. The sensory and process precisions $\pi_{\tilde{z}}, \pi_{\tilde{w}}$ are fixed, to show here only the basic disturbance rejection property of PID controllers (Åström, 1995; Sontag, 2003). In Fig. 6.2a, after the car is exposed to some new external condition (e.g., wind) represented in Fig. 6.2c and not encoded in the controller's generative model, the regulation process brings the velocity of the car back to the desired state after a short transition period. Fig. 6.2b shows how sudden changes in the acceleration of the car are quickly cancelled out in accord with the specified prior $\eta'_x = 0$ km/h^2. The action of the car is then shown, as one would expect (Sontag, 2003), to counteract the external force v, Fig. 6.2c.

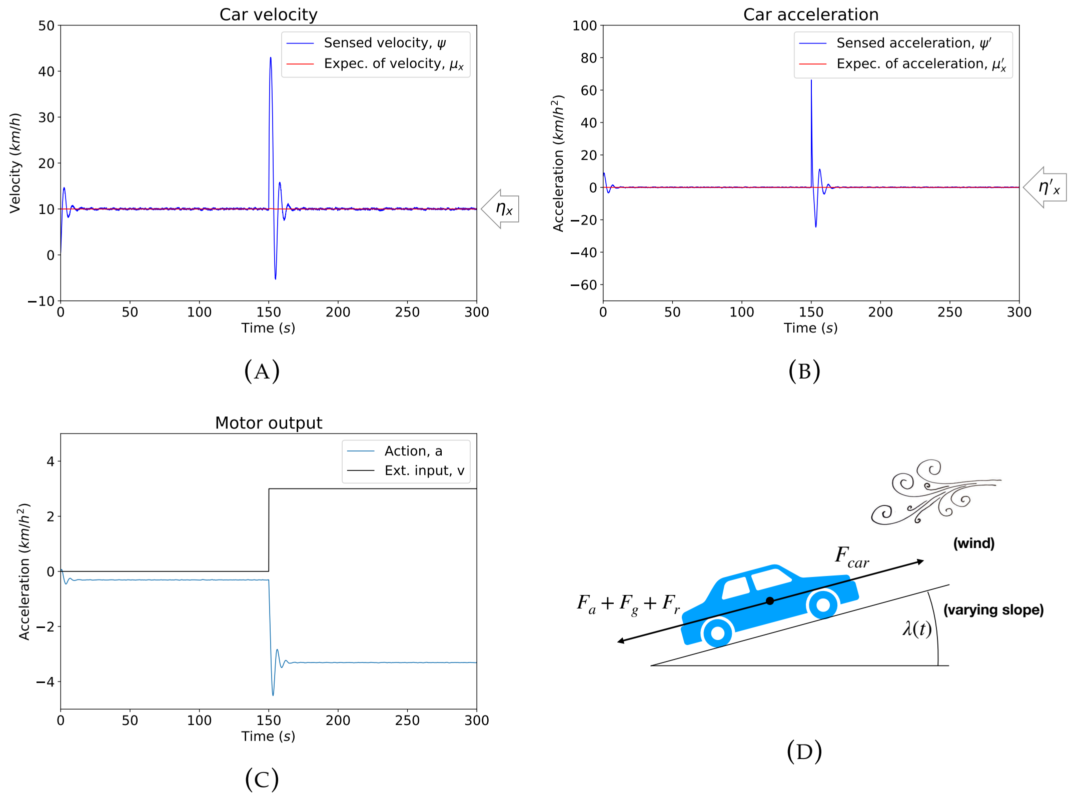


FIGURE 6.2: **A cruise controller based on PI control under active inference.** (a) The response of the car velocity over time with a target state, or prior in our formulation, $\eta_x = 10$ km/h, $\eta_{x'} = 0$ km/h^2. (b) The acceleration of the car over time with a specified prior $\eta'_x = 0$ km/h^2. (c) The external force v, introduced at $t = 150$s, models a sudden change in the environmental conditions, for instance wind or change in slope. Action obtained via the minimisation of variational free energy with respect to a and counteracts the effects of v. The motor action is never zero since we assume a constant slope, $\lambda = 4°$ (see table 6.1, section 6.3). (d) The model car we implemented, where v could be thought as a sudden wind or a changing slope.

6.3.1 Responses to external and internal changes

It is often desirable for a PID regulator to provide different responses to external perturbations (e.g., wind), which should be rather rapid, and to internal updates (e.g., a shift in target velocity) which should be relatively smooth (Åström, 1995; Åström and Hägglund, 2004), see also section 6.1.1. It is not, however, trivial to identify and isolate parameters that contribute to these effects (Araki and Taguchi, 2003; Ang, Chong, and Li, 2005; Johnson and Moradi, 2005), and thus to tune these properties independently. It has been suggested that in order to achieve such decoupling, a controller with two degrees of freedom is necessary (Araki and Taguchi, 2003; Åström and Hägglund, 2004). Such controller can be thought to contain a feedforward model of the dynamics of the observed/regulated system (Åström and Murray, 2010). In our implementation, this is elegantly achieved by construction, since active inference is based on generative (forward) models. Specifically, we can fix the response to external forces by setting the expected sensory precisions $\pi_{\tilde{z}}$ (i.e., PI gains) but then independently tune the response to changes in the setpoint by altering the expected process precisions $\pi_{\tilde{w}}$ on the priors, see Fig. 6.3 and Fig. 6.4.

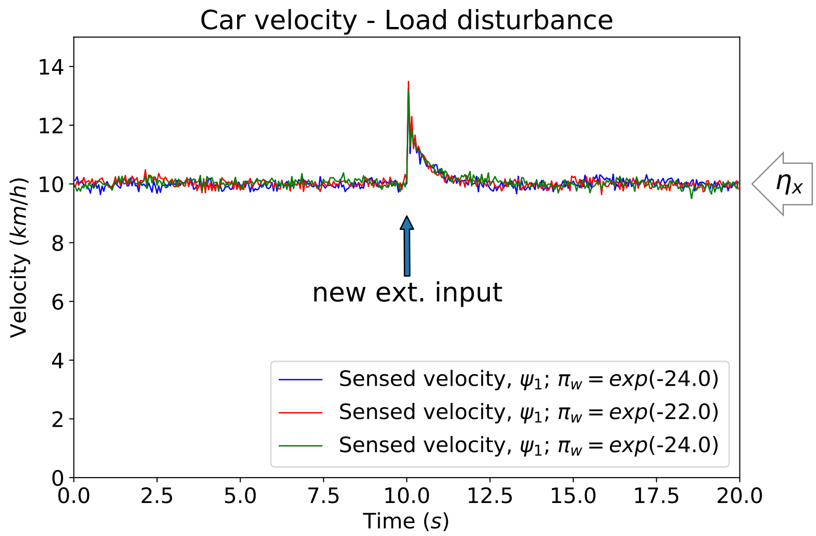


FIGURE 6.3: **Responses to load disturbances.** The same load disturbance ($v = 3.0km/h^2$) is applied with varying expected process precisions $\pi_{\tilde{w}}$ where $\pi_w = \{\exp(-24), \exp(-22), \exp(-20)\}$. Expected sensory log-precisions $\pi_{\tilde{z}}$ are fixed over the duration of the simulations, with $\mu_{\gamma_z} = 1$. The simulations were 300s, with an external disturbance introduced at $t = 150$s. Here we report only a 20 seconds time window around the change in conditions.

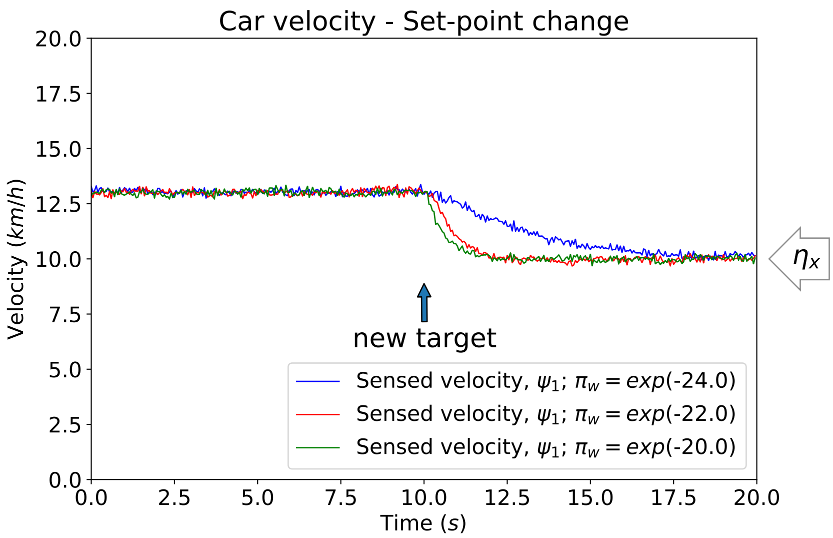


FIGURE 6.4: **Responses to load set-point changes.** A similar example for changes in the target velocity of the car, from $\eta_x = 13km/h$ to $\eta_x = 10km/h$, tested on varying expected process precisions $\pi_{\tilde{w}}$ where $\pi_w = \{\exp(-24), \exp(-22), \exp(-20)\}$. The simulations were 300s, with a different target velocity introduced at $t = 150$s. Here we report only a 20 seconds time window around the change in conditions.

In the limit for process prediction errors $\pi_{\tilde{w}}(\tilde{\mu}'_x + \alpha(\tilde{\mu}_x - \tilde{\eta}_x))$ much larger than the sensory ones $\pi_{\tilde{z}}(\tilde{\psi} - \tilde{\mu}_x)$ and with fixed expected sensory precisions $\pi_{\tilde{z}}$, the response to load disturbances is invariant (Fig. 6.3). A new target velocity for the car creates different responses with varying $\pi_w = \{\exp(-24), \exp(-22), \exp(-20)\}$[2]. Larger $\pi_{\tilde{w}}$ values imply an expected low uncertainty on the dynamics (i.e., changes to the set-point are not encoded and therefore not expected) and are met almost instantaneously with an update of expected hidden states $\tilde{\mu}_x$, matched by suitable actions a. On the other hand, smaller $\pi_{\tilde{w}}$ account for higher variance/uncertainty and thus changes in the target velocity are to be expected, making the transitions to new reference values slower, as seen in Fig. 6.4.

6.3.2 Optimal tuning of PID gains

One of the main goals of modern design principles for PID controllers is to find appropriate tuning rules for the gains on the prediction errors: proportional, integral and derivative terms. Existing approaches are often limited (Åström, Panagopoulos, and Hägglund, 1998; Åström and Hägglund, 2001; Johnson and Moradi, 2005; Ang, Chong, and Li, 2005; Åström and Hägglund, 2006) and a unifying theoretical framework guiding how the gains should be optimally tuned is elusive (Åström and Hägglund, 2001; Ang, Chong, and Li, 2005). In general, the proportional term must bring a system to the target state in the first place, the integral of the error should promptly deal with errors generated by steady state inputs not accounted by a model (Sontag, 2003), while the derivative term should reduce the fluctuations by controlling changes in the derivative of a variable (Åström and Murray, 2010). In our car example, this could mean for example controlling the velocity of the vehicle in spite of changes such as the presence of wind or variations in slope of the road (I term) and avoiding unnecessary changes in accelerations close to the target (D term, often not used for cruise control problems anyway (Åström and Murray, 2010)). In our treatment of PID controllers as approximate Bayesian inference, the controllers' gains k_i, k_p, k_d become equivalent to sensory precisions $\pi_z, \pi_{z'}, \pi_{z''}$, cf. equation (6.13) and equation (6.14). Following Friston, Trujillo-Barreto, and Daunizeau (2008), Friston (2008a), and Friston et al. (2010b), we thus propose to optimise these precisions to minimise the path integral of variational free energy (or free action), assuming that parameters and hyperparameters change on a much slower time scale. To do so, we extend our previous formulation and replace fixed sensory precisions $\pi_z, \pi_{z'}, \pi_{z''}$ with *expected* sensory precisions $\mu_{\pi_z}, \mu_{\pi_{z'}}, \mu_{\pi_{z''}}$, derived from a Laplace approximation applied not only to hidden states x but extended also to these hyperparameters, now considered as random variables to be estimated, rather than fixed quantities (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston et al., 2010b), see Chapter 3.

Active inference provides then an analytical criterion for the tuning of PID gains in the temporal domain, where otherwise mostly empirical methods or complex

[2]Precisions on higher embedding orders are built, in both cases, using a smoothness (i.e., decay) factor of $1/2$, see Friston (2008a)

methods in the frequency domain have insofar been proposed (Rivera, Morari, and Skogestad, 1986; Åström, 1995; Åström, Panagopoulos, and Hägglund, 1998; Åström and Hägglund, 2006). In frameworks used to implement active inference, such as DEM (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston, 2008a), parameters and hyperparameters are usually assumed to be conditionally independent of hidden states based on a strict separation of time scales (i.e., a mean-field approximation). This assumption prescribes a minimisation scheme with respect to the path-integral of free energy, or free action, requiring the explicit integration of this functional over time. In our work, however, for the purposes of building an online self-tuning controller, we will treat expected sensory precisions as conditionally dependent but changing on a much slower time-scale with respect to states x, using a second order online update scheme based on generalised filtering (Friston et al., 2010b), as explained in Chapter 3. The controller gains, $\mu_{\pi_z}, \mu_{\pi_{z'}}, \mu_{\pi_{z''}}$, will thus be updated following a reparametrisation that ensures they are positive (precisions are defined as non negative and control is achieved only with negative feedback), following the scheme presented in Chapter 3, equation (3.41).

For practical purposes, the second order system presented in equation (3.41) is usually reduced to a simpler set of first order differential equations (Buckley et al., 2017):

$$\begin{aligned} \dot{\mu}_{\gamma_{\tilde{z}}} =& \tilde{\phi} \\ \dot{\tilde{\phi}} =& -\frac{\partial F}{\partial \mu_{\gamma_{\tilde{z}}}} - \kappa \tilde{\phi} \end{aligned} \tag{6.17}$$

where $\tilde{\phi}$ is a prior on the motion (i.e., only the rate of change in this case) of hyperparameters γ which encodes a "damping" term for the minimisation of free energy F. This term enforces hyperparameters to converge to a solution close to the real steady state thanks to a drag term for $\kappa > 0$[3]. The parametrisation of expected precisions in terms of log-precisions $\gamma_{\tilde{z}}$, in fact, makes the derivative of the free energy with respect to log-precisions strictly positive ($\frac{\partial F}{\partial \gamma_{\tilde{z}}} > 0$), not providing a steady-state solution for the gradient descent (Friston et al., 2010b). This "damping" term stabilises the solution, reducing the inevitable oscillations around the real equilibrium of the system. Given the free energy defined in equation (6.6), with $\exp(\mu_{\gamma_{\tilde{z}}})$ replacing $\pi_{\tilde{z}}$, the minimisation of expected sensory log-precisions (or "log- PID gains") is

[3] $\kappa = 5$ in our simulations

prescribed by the following equations:

$$\begin{aligned}
\dot{\mu}_{\gamma_z} =& \phi_z \\
\dot{\phi}_z =& -\frac{\partial F}{\partial \mu_{\gamma_z}} - \kappa \phi_z = -\frac{1}{2}\Big[\exp\left(\mu_{\gamma_z}\right)(\psi - \mu_x)^2 - 1\Big] - \kappa \phi_z \\
\dot{\mu}_{\gamma_{z'}} =& \phi_{z'} \\
\dot{\phi}_{z'} =& -\frac{\partial F}{\partial \mu_{\gamma_{z'}}} - \kappa \phi_{z'} = -\frac{1}{2}\Big[\exp\left(\mu_{\gamma_{z'}}\right)(\psi' - \mu_x')^2 - 1\Big] - \kappa \phi_{z'} \\
\dot{\mu}_{\gamma_{z''}} =& \phi_{z''} \\
\dot{\phi}_{z''} =& -\frac{\partial F}{\partial \mu_{\gamma_{z''}}} - \kappa \phi_{z''} = -\frac{1}{2}\Big[\exp\left(\mu_{\gamma_{z''}}\right)(\psi'' - \mu_x'')^2 - 1\Big] - \kappa \phi_{z''}
\end{aligned} \tag{6.18}$$

This scheme introduces a new mechanism for the tuning of the gains of a PID controller, allowing the controller to adapt to adverse and unexpected conditions in an optimal way, in order to avoid oscillations around the target state.

In Fig. 6.5 the controller for the car velocity is initialised with suboptimal sensory log-precisions $\mu_{\gamma_{\tilde{z}}}$, i.e., log-PI gains. The parameters were initially not updated (Fig. 6.5d) to allow the controller to settle around the desired state, see Fig. 6.5a. The adaptation begins at $t = 30$s and is stopped at $t = 150$s, when an external force is introduced, to test the response of the controller after the gains have been optimised. With the adaptation process, the controller becomes more responsive when facing external disturbances (cf. Fig. 6.2), quickly and effectively counteracted by prompt changes in controls, see Fig. 6.5c. The optimisation of the gains through $\mu_{\gamma_{\tilde{z}}}$ without extra constraints (if not the stopping condition we imposed at $t = 150$s, after the adaptation reaches a steady-state) effectively introduces an extremely responsive controller: cancelling out the effects of unwanted external inputs, such as wind in our cruise control example, but also more sensitive to measurement noise. In Fig. 6.6 we show summary statistics with the results of the adaptation of the gains. Following the examples in Fig. 6.2 and Fig. 6.5, we simulated 20 different cars with expected sensory log-precisions $\mu_{\gamma_{\tilde{z}}}$ sampled uniformly in the interval $[-4, -2]$ and expected process log-precisions $\mu_{\gamma_{\tilde{w}}}$ in the interval $[-23, -21]$. We initially maintained (i.e., no adaptation) the same hyperparameters and introduced a load disturbance at $t = 150$s, then repeated the simulations (20 cars) with the same initial conditions allowing for the adaptation of expected sensory log-precisions as log-PI gains after $t = 30$s, as in Fig. 6.5. Following (Hägglund, 1995), we measured the performance of the controllers by defining the integral absolute error (IAE):

$$IAE = \int_t^{t+\tau} |e(t)| \;\mathrm{d}t \tag{6.19}$$

between two zero-crossings: the last time the velocity was at the target value before a disturbance is introduced, assumed to be $t = 150$ in our case, and the first time the velocity goes back to the target after a disturbance is introduced ($t + \tau$). To compute $t+\tau$, we took into account the stochasticity of the system and errors due to numerical

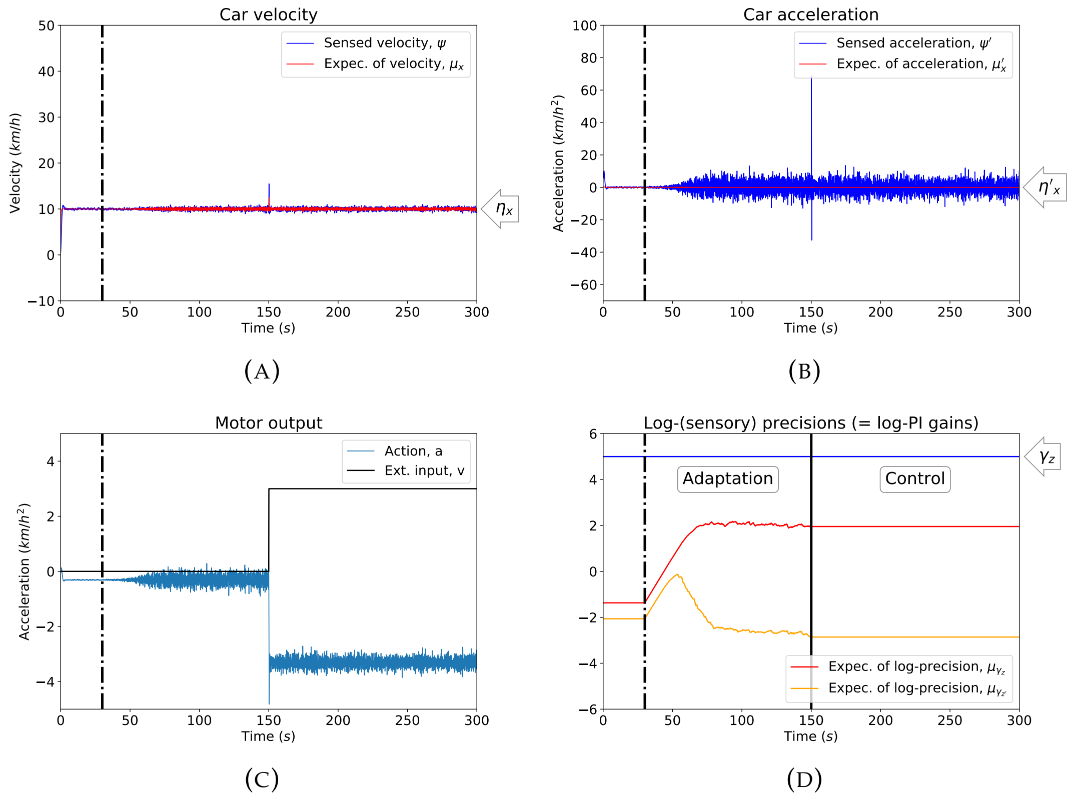


FIGURE 6.5: **Optimising PID gains as expected sensory log-precisions** $\mu_{\gamma_{\tilde{z}}}$. This example shows the control of the car velocity before and after the optimisation of $\mu_{\gamma_{\tilde{z}}}$ (before and after the vertical dash dot black line) is introduced. (a) The velocity of the car. The controller is more responsive to sudden external forces, e.g., wind, immediately regulated against (cf. Fig. 6.2a), but as a trade-off, the variance of the velocity is also increased. (b) The acceleration of the car. (c) The action of the car, with an external disturbance introduced at $t = 150$s. (d) The optimisation of expected sensory precisions $\mu_{\gamma_{\tilde{z}}}$ and their convergence to an equilibrium state, after which the optimisation is stopped before introducing an external force. The blue line represents the true log-precision of observation noise in the system, $\gamma_z = \gamma_{z'} = 5$.

approximations, considering the case for the real velocity to be within a ± 0.5 km/h interval away from the target value. The IAE captures the impact of oscillations on the regulation problem by integrating the error over the temporal interval where the car is pushed away from its target due to some disturbance (for more general discussions on its role and uses see Åström (1995)). As we can see in Fig. 6.6, the IAE converges to a single value for all cars (taking into account our approximation of a ± 0.5 km/h interval while measuring it) and is clearly lower when the adaptation mechanism for expected sensory log-precisions is introduced, making the controller very responsive to external forces and thus reducing the time away from the target velocity, see Fig. 6.5 for an example.

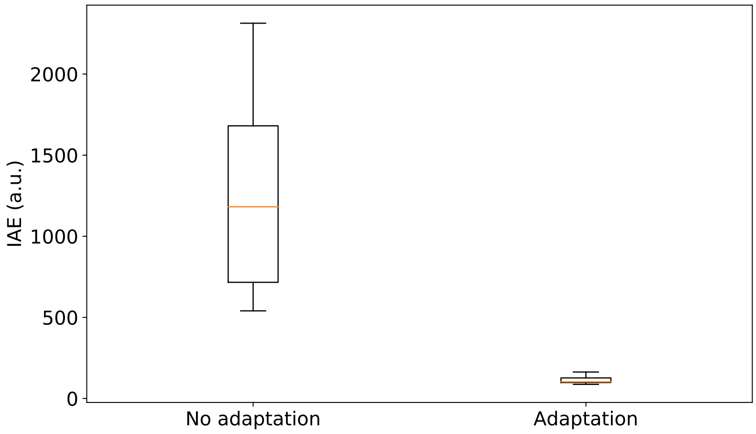


FIGURE 6.6: **Performance of PID controllers with and without adaptation of the gains based on the minimisation of free energy.** The integral absolute error (IAE) is used to measure the effects of the oscillations introduced by a single load disturbance at $t = 150$s (see text for the exact definition of the IAE).

6.4 Measurement noise and model uncertainty in active inference

Nowadays, it is common to include two more desiderata for the design of PID controllers (see section 6.1 and Åström and Hägglund (2001)) in order to characterise and tune their response to (1) different types of measurement noise and (2) their robustness to model uncertainty, inherent in simple approximate controllers (Åström and Hägglund, 2001; Åström and Hägglund, 2006). In our example, these properties map, respectively, to the response of a car given time-varying noise and to the available knowledge of a system, e.g., the working range of a controller or the type of disturbances affecting the car.

In particular, the former describes the behaviour of a PID controller in presence of noise on the observed variables by modulating the decay of different prediction errors in equation (6.13). It is known that this response is (in the limit for $t \to \infty$ and with the assumption of a system at equilibrium) inversely proportional to the integral gain (Åström, 1995; Åström and Hägglund, 2006). In our case however, we have a more general and trivial relationship where the integral gain k_i is, by construction, equivalent to the inverse variance (i.e., precision) of the measurement noise π_z, see equation (6.13) and equation (6.14). The remaining gains k_p, k_d can then be seen as encoding the uncertainty (i.e., precision) of higher orders of motion when the measurement noise is effectively coloured, otherwise just approximating possible correlations of the observed data over time.

On the other hand, the robustness to model uncertainty can be seen in terms of expected process log-precisions $\mu_{\gamma_{\tilde{w}}}$ encoding (again by construction) the amplitude of fluctuations due to unknown effects on the dynamics (Friston, 2008a). By modulating the prior dynamics of a system, these hyperparameters assume then a double role, they can either: (1) passively describe (estimate) the dynamics of a system (cf. Kalman filters (Kalman, 1960b)) or (2) actively impose desired trajectories on observations that can be implemented through actions on the world, as explained in section 6.2 and more generally in Chapter 4. With these conditions at the extremes, a spectrum of intermediate behaviours is also possible, with $\mu_{\gamma_{\tilde{w}}}$ enacting different sensorimotor couplings by weighting the importance of objective information and desired states/priors of a system.

In the majority of the formulations of control problems, the properties of measurement noise and model uncertainty (especially their (co)variance) are assumed to be constant over time. Often, these parameters need also to be adapted to different systems since their properties are likely to be different. In section 6.3.2, we proposed an optimisation method for the gains of a PID controller based on active inference that here we exploit for time changing properties of the noise of a system, and that we show in an example when the measurement noise suddenly increases. In our car example, we could think of a drop in performance of the sensors recording velocity.

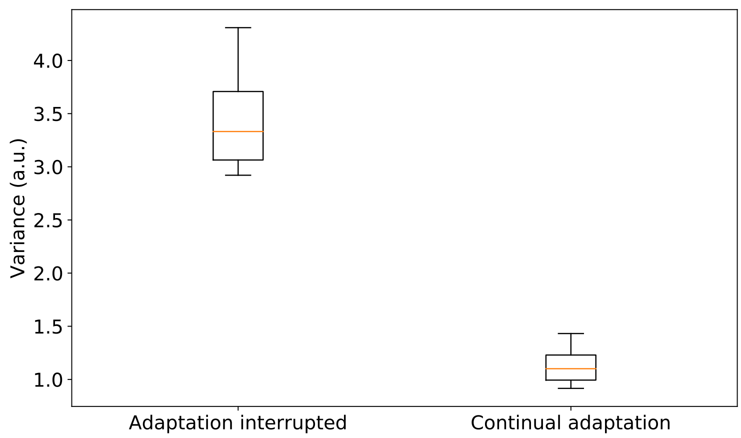


FIGURE 6.7: **Performance of PID controllers with a sudden increase in measurement noise.** 20 cars simulated in the case where measurement noise is increased at $t = 150$s during the 300s simulations. We report aggregate results with the variance from the target value measured over the last 25% $(225 < t < 300\text{s})$ of a simulation. We show (1) the case for adaptation of the gains of the PI controller (through expected sensory log-precisions, or log-PI gains, $\mu_{\gamma_{\tilde{z}}}$) interrupted before the measurement noise drastically changes, and (2) the case where the adaptation process persists for the entire duration of the simulations.

We simulated 20 cars for 300s with adaptation of expected sensory log-precisions (or log-PI gains) $\mu_{\gamma_{\tilde{z}}}$, initiated at $t = 30$s and stopped at $t = 150$s. At $t = 150$s we then decreased the log-precision of measurement noise (n.b. not the expectation on the log-precision) from $\gamma_z = 5$ to $\gamma_z = 2$ for the rest of the simulations and stopped the adaptation process. We then had 20 cars with the same set up but for which adaptation was not halted after the increased measurement noise. To represent the difference, we measured the variance of the real velocity of the cars (without measurement noise to avoid biases), from $t = 225$s to $t = 300$s to allow the velocity to settle after the transient due to the sudden change. Controllers continuously adapting their gains are shown to be more robust to persistent changes in noise.

In the case of model uncertainty, given the dual role of $\mu_{\gamma_{\tilde{w}}}$ explained above, i.e., encoding prior dynamics reflecting both real properties of the environment and desired trajectories imposed on the system to regulate, it is harder to show the update of expected precisions without compromising the control of the car. The optimisation we proposed is, in fact, not intrinsically biased towards the control of a system, i.e., we externally imposed that as a condition for the agent. While having more flexible priors, an agent could potentially begin to account for uncertainty in the world rather than forcibly change its observations to reach its target.

6.5 Discussion

In this chapter we developed a minimal account of regulation and control mechanisms in biological systems based on active inference, a process theory for perception, action and higher order functions expressed via the minimisation of variational free energy (Friston, Kilner, and Harrison, 2006; Friston et al., 2010a; Friston, 2010b; Buckley et al., 2017). Our implementation constitutes an example of the parsimonious, action-oriented models described in Clark (2015a) and Clark (2015b), connecting them to methods from classic control theory. We focused in particular on Proportional-Integral-Derivative (PID) control, both extensively used in industry (Åström, 1995; Ang, Chong, and Li, 2005; Johnson and Moradi, 2005; Åström and Hägglund, 2006) and more recently emerging as a model of robust feedback mechanisms in biology, implemented for instance by bacteria (Yi et al., 2000), amoeba (Yang and Iglesias, 2006) and gene networks (Ang et al., 2010), and in psychology (Ritz et al., 2018). PID controllers are ubiquitous in engineering mostly due to the fact that one needs only little knowledge of the process to regulate. In the biological sciences, this mechanism is thought to be easily implemented even at a molecular level (Chevalier et al., 2018) and to constitute a possible account for limited knowledge of the external world in simple agents (Sontag, 2003).

Following our previous work on minimal generative models (see Chapter 5), we showed that this mechanism corresponds, in active inference terms, to linear generative models for agents that only approximate properties of the world dynamics. Specifically, our model describes linear dynamics for a single hidden or latent state and a linear mapping from the hidden state to an observed variable, representing knowledge of the world that is potentially far removed from the real complexity behind observations and their hidden variables. To implement such model, we defined a generative model that only approximates the environment of an agent and showed how under a set of assumptions including analytic (i.e., non-Markovian, differentiable) Gaussian noise and linear dynamics, this recapitulates PID control. A crucial component of our formulation is the presence of low sensory precision parameters on proprioceptive prediction errors of our free energy function or equivalently, high expected variance of proprioceptive signals. These low precisions play two roles during the minimisation of free energy: (1) they implement control signals as predictions of proprioceptive input influenced by strong priors (i.e., desires) rather than by observations, see equation (6.11) and Friston et al. (2010a), and (2) they reflect a belief that there are large exogenous fluctuations (low precision = high variance) in the observed proprioceptive input. This last point can be seen as the well known property of the Integral term (Åström and Murray, 2010; Sontag, 2003) of PID controllers, dealing with unexpected external input (i.e., large exogenous fluctuations). The model represented by derivatives $\partial\tilde{\psi}/\partial a$ encodes then how actions a approximately affect observed proprioceptive sensations $\tilde{\psi}$, with an agent implementing a sensorimotor mapping that does not match the real dynamics of actions applied to the environment. The formulation in equation (6.7) and equation (6.8) can in general be applied to different tasks, in the same way PID control is used in different problems without specific knowledge of the system to regulate.

The generative model we used is expressed in generalised coordinates of motion, a mathematical construct used to build non-Markovian continuous stochastic models. Their importance has been expressed before (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston, 2008a; Friston et al., 2010b), for the treatment of real world processes best approximated by continuous models and for which Markov assumptions don't really hold (see also Kim (2018) for discussion). The definition of a *generalised* state-space model provides then a series of weighted prediction errors and their higher orders of motion from the start, with PID control emerging as the consequence of an agent trying to impose its desired prior dynamics on the world via the approximate control of its observations on different embedding orders (for I, P and D terms). In this light, the ubiquitous efficacy of PID control may thus reflect the fact that the simplest models of controlled dynamics are first-order approximations to generalised motion. This simplicity is mandated because the minimisation of free energy is equivalent to the maximisation of model evidence, which can be expressed as accuracy minus complexity (Friston, 2010b; Clark, 2015b). On this view, PID control emerges via the implementation of constrained (parsimonious, minimum complexity) generative models that are, under some constraints, the most effective (maximum accuracy) for a task.

In the control theory literature, many tuning rules for PID gains have been proposed (e.g., Ziegler-Nichols, IMC, etc., see (Åström, 1995; Åström and Hägglund, 2006) for a review) and used in different applications (Åström, 1995; Åström, Panagopoulos, and Hägglund, 1998; Ang, Chong, and Li, 2005; Johnson and Moradi, 2005;

Åström and Hägglund, 2006), however most of them produce quite different results, highlighting their inherent fit to only one of many different goals of the control problem. It is also unclear what the criterion, or set of criteria, for optimality of PID controllers should be: minimum variance, minimum time, stability, robustness etc.. With our active inference formulation, we argue that different criteria can and should be expressed within the same set of equations in order to better understand their implications for a system. Modern approaches to the study of PID controllers propose four points as fundamental features to be considered for the design of a controller (Åström and Hägglund, 2001):

- load disturbance response
- set-point response
- measurement noise response
- robustness to model uncertainty.

In our formulation, these criteria can be interpreted using precision (inverse variance) parameters of different prediction errors in the variational free energy, expressing the the uncertainty associated to observations and priors, as reported in table 6.2, see also section 6.4 for further reference.

TABLE 6.2: **Active inference as a general framework for PID controllers.**

Criterion	**Mapped to**	**Advantages in active inference**
Load disturbance response	$\mu_{\pi_{\tilde{z}}}$	Intuitively expressed via the expected inverse variance of the observations (i.e., precision), with low variance implying a fast response and vice versa (see section 6.3.1 and section 6.3.2)
Set-point change response	$\mu_{\pi_{\tilde{w}}}$	Natural formulation of PID controllers with two degrees of freedom derived from sensory and process precisions and expressed as a Bayesian inference process (see section 6.3.1)
Measurement noise response	$\mu_{\pi_{\tilde{z}}}$	Straightforward interpretation of PID gains as (expected) inverse variances of different embedding orders of measurement noise (see section 6.4)
Robustness to model uncertainty	$\mu_{\pi_{\tilde{w}}}$	Direct mapping of model uncertainty to expected variances of the fluctuations, representing unknown dynamics, of the system to control (see section 6.4)

After establishing the equivalence between PID control and linear approximations of generalised motion in generative models, we showed that the controllers' gains, k_i, k_p, k_d, are in our formulation equivalent to expected precisions, $\mu_{\pi_z}, \mu_{\pi_{z'}}, \mu_{\pi_{z''}}$, for which a minimisation scheme is provided in Friston, Trujillo-Barreto, and Daunizeau (2008), Friston (2008a), and Friston et al. (2010b). The basic version of this optimisation produces also promising results in presence of time-varying measurement (white) noise in the simulated car (see Fig. 6.7 in section 6.4). If the adaptation is halted on a system with fixed measurement noise, it can be used to effectively deal with load disturbances, external forces acting against a system reaching his target (see Fig. 6.5), e.g., a change in chemicals concentration for a bacterium.

Future extensions could provide a more principled way of dealing with these two (and possibly other) conflicting cases, an issue that can be solved by introducing suitable hyperpriors (priors on hyperparameters) expressing the confidence of a system regarding changes in measurement noise via the use of precisions on hyperpriors (Friston, 2008a). High confidence (i.e., high precision on hyperpriors) would imply that a system should quickly react to sudden changes, both in measurement noise and other disturbances, since they are unexpected. On the other hand, low confidence (i.e., low precision on hyperpriors) would make a system's reaction to new conditions slower since such changes are expected. A trade-off between these conditions, with appropriate knowledge of a system or a class of systems and introduced in the form of hyperpriors, would then make the process completely automatised, taking advantage of empirical Bayes for learning such hyperpriors (Friston, 2010b). By extending our proposition with priors on precisions (i.e., hyperpriors) we can also, in principle, cast more criteria for the controller, expressing different requirements for more complex regulation processes. Given the fact that any optimality criterion can be recast as a prior, following the complete class theorem (Ferguson, 1967; Brown, 1981), as long as we know how to represent these rules as priors for the controller, we can provide any combination of requirements and tune the parameters in a straightforward fashion.

6.6 Conclusion

PID controllers are robust controllers used as a model of regulation for noisy and non-stationary processes in different engineering fields (Åström and Hägglund, 2006; Åström and Murray, 2010). More recently, they have also been proposed as behavioural models of adaptive learning in humans (Ritz et al., 2018) and as mechanistic explanations of different functions of systems in microbiology (Yi et al., 2000; Yang and Iglesias, 2006; Ang et al., 2010). Their utmost relevance to the natural sciences is becoming clear, with implementations now proposed at the level of simple biomolecular interactions (Briat, Gupta, and Khammash, 2016; Chevalier et al., 2018). PID controllers are renowned for their simplicity and straightforward interpretation, however they do not guarantee optimality, so an interpretation of this control strategy in terms of more general mathematical principles is still missing.

Active inference has been proposed as a general mathematical theory of life and cognition according to the minimisation of variational free energy (Friston, 2010b). On this view, biological agents are to be seen as homeostatic systems maintaining their existence via the the minimisation of free energy. This process is implemented via the estimation and prediction of latent variables in the world (equivalent to perception) and the control of sensory inputs with behaviours accommodating normative constraints of an agent. It is often described as an extension of optimal control theory with deep connections to Bayesian inference (Friston, 2011). While methods such as PID control are still widely adopted as models of biological systems, it is unclear how general theories such as active inference connect to practical implementation of homeostatic principles such as PID control. In this chapter we proposed a way to connect these two perspectives showing how PID controllers can be seen as a special case of active inference. This account is based on the definition of a linear generative model for an agent approximating the dynamics of its environment, potentially very different from the information represented by the model. The model is expressed in generalised coordinates of motion (Friston, 2008a; Buckley et al., 2017; Kim, 2018) with prediction errors at different temporal orders for integral, proportional and derivative components emerging naturally as a consequence of an agent assuming non-Markovian dynamics on its sensory input. Through the active inference we also proposed the implementation of a mechanism for the optimisation of the gains of a PID controller, i.e., the weights of different prediction errors, now interpreted as precision parameters encoding the uncertainty of different variables from the perspective of an agent.

In this chapter I introduced another example of an agent described by a generative model whose formulation is deeply divorced from the dynamics of its milieu, reinforcing the idea that Bayesian accounts of cognition need not consider agents as mirrors of their world. Furthermore, following some of the claims regarding the FEP as a possible unifying theory of *all* biologyical systems (Friston, 2012; Friston, 2013), this work constitutes also a direct connection between the FEP and models of behaviour in simple (down to unicellular) biological organisms, e.g., (Yi et al., 2000).

Chapter 7

Modularity, the separation principle and active inference

The assumption that action and perception can be investigated independently is entrenched in theories, models and experimental approaches across the brain and mind sciences. In cognitive science, this has been a central point of contention between computational and 4E (enactive, embodied, extended and embedded) theories of cognition, with the former embracing the “classical sandwich", modular, architecture of the mind and the latter advocating a more holistic view of cognitive functions as depending on agent-environment dynamics. In this chapter, we suggest that the modularity of action and perception at the core of computational theories may be seen in analogy with the *separation principle* of control theory. Furthermore, we argue that this principle provides formal criteria that can be used to evaluate the implications of the modularity of action and perception in the cognitive and natural sciences. We also claim that real-time responses to feedback from the environment, one of the main ideas proposed by 4E approaches in contrast to more traditional proposals, can still be consistent with cognitivist theories, again in analogy with the separation principle. Finally we argue that an emerging theory in the cognitive and brain sciences, active inference, by extending ideas derived from control theory to the study of biological systems, disposes of the separation principle. In doing so, it describes non-modular models of behaviour strongly aligned with some aspects of non-reductionist theories of cognition.

This chapter provides one of the strongest arguments for the connection between the FEP and 4E cognition. By using its roots in control theory, I will argue that the FEP is in conflict with the dominant view of perception and action as separate processes adopted by the vast majority of models in psychology and neuroscience. Following the literature reviewed in Chapter 2, I will first introduce the core mathematical formulations of such models based on the separation principle and its most well known instantiation, the LQG framework. I will then focus on the differences with the active inference proposal, implementing a standard example (the double integrator) classically used for the study of LQG systems, comparing the LQG solution of the problem with an active inference one. As in Chapter 4, the generative model of the agent and the generative process of the environment will be relatively similar since the problem's formulation is extremely simple to begin with. Important differences will however emerge due to the priors encoded by our active inference agent, "hallucinating" (imaginary) springs and pistons attached to its body that fulfil its desires. The chapter is based on material from Baltieri and Buckley (2018b), Baltieri and Buckley (2019b), and Baltieri and Buckley (2019a).

7.1 Background

Can perception and action be studied as separate processes?

In cognitive science, the hypothesis that the mind is modular originated with Fodor's work (Fodor, 1983; Coltheart, 1999; Barrett and Kurzban, 2006; Prinz, 2006), formalising the idea that the perceptual and motor systems should be considered as formed by separate and informationally encapsulated components, sitting at the periphery of an organism (for a more historical review see Boden (2006)). This view has inspired decades of research in different areas of cognition that have recently been described by the so-called "classical sandwich" architecture of cognitive systems, whereby cognition sits in between perception and action, effectively rendering them almost autonomous (Hurley, 2001; Wilson and Foglia, 2017). This view contrasts with non-reductionist theories of the mind such as 4E (enactive, embodied, embedded and extended), suggesting that situated, dynamical interactions with the environment are critical to explain cognitive processes. In doing so, 4E proposals reject the hypothesis of segregated perceptual and motor components (Van Gelder, 1995; Clark, 1998; Wilson, 2002; Beer and Williams, 2015; Di Paolo, Buhrmann, and Barandiaran, 2017), now seen as integrated and strongly coupled by feedback mechanisms mediated by the environment.

In this chapter we argue, however, that the emphasis of feedback is not enough to distinguish 4E theories from the ones related to the classical sandwich architecture (Varela, Thompson, and Rosch, 1991; Beer, 1995). We will see that the modular view can (often implicitly) survive in modern studies of action and perception even in the presence of closed sensorimotor loops. To ground this argument and related discussions on Fodorian (Fodor, 1983; Coltheart, 1999; Prinz, 2006) and post-Fodorian (Barrett and Kurzban, 2006) modularity, we make use of formal frameworks that have emerged from information and control theory, already widely exploited in modern theories of perception and action based on processes of estimation/inference and control (Knill and Richards, 1996; Rao and Ballard, 1999; Kawato, 1999; Wolpert and Ghahramani, 2000; Todorov, 2004; Friston, 2010b; Friston, 2011; Adams, Shipp, and Friston, 2013) (see Chapter 2). Perception, on this view, is modelled as a process of estimation of the hidden or latent variables of the world given noisy and often inaccurate observations. Action, on the other hand, is accounted for with theories of optimal control, suggesting possible optimality principles for the implementation of motor actions and behaviour more in general. In this light, we then argue that the so-called *separation principle* of estimation and control in control theory provides concrete grounding of Fodor's modularity with regards to action and perception. After presenting this principle and its connections to the idea of modularity, we will discuss its role in the study of cognitive and natural systems. We finally propose active inference (Friston et al., 2010a; Friston, 2010b) as an alternative view that, we argue, openly rejects the separation principle, thus supporting non-modular 4E arguments while maintaining its control/information theoretical roots consistent with modern approaches to action and perception.

7.2 Modularity as an analogy of the separation principle

The *separation principle* of control theory provides a set of necessary and sufficient conditions under which an optimal controller can be constructed by combining a Kalman(-Bucy) filter and a Linear Quadratic Regulator that can be treated independently (Wonham, 1968; Åström, 1970; Anderson and Moore, 1990; Stengel, 1994; Georgiou and Lindquist, 2013). This methodology is widely adopted in control theory and can be used to build controllers for noisy and uncertain systems where environmental states are only partially observable. Separating estimation and control is practically desirable because it becomes then possible to *optimally* solve the estimation problem and subsequently use the output estimate to build an *optimal* controller. The idea of the separation theorem is also very closely related to the certainty equivalence principle described in econometrics and decision making (Simon, 1956; Theil, 1957; Bar-Shalom and Tse, 1974; Åström, 1970; Stengel, 1994). In information theory, Shannon (1948) also previously introduced a definition of a separation principle, to explain coding via two (separate) phases of source compression and channel coding (Gastpar, Rimoldi, and Vetterli, 2003). The connections between this notion and the separation principle in control theory have become more clear in recent years, thanks to work showing how Shannon's description captures and potentially generalises the results from control theory, see for instance Tatikonda (2000), Tanaka, Esfahani, and Mitter (2015), and Fox and Tishby (2016). In this thesis, the focus will however be on the principle traditionally described in control theory.

The separation principle rests on a set of assumptions that, following (Åström, 1970; Anderson and Moore, 1990; Stengel, 1994), can be summarised as:

1. linear process dynamics and observation laws describing the environment and its latent variables

2. Gaussian white noise in both process and measurement equations/laws

3. known (co)variance matrices representing uncertainty of both process and measurement noises

4. a quadratic cost function used to measure the performance of a system

5. known inputs for the estimator (i.e., the Kalman-Bucy filter), since the estimator needs to have access to all the variables and parameters, external and internal ones, affecting the inference of hidden states.

In the next sections we will first briefly introduce Kalman(-Bucy) filters and Linear Quadratic Regulator (LQR) as the optimal estimator and controller, respectively, for the linear case. Linear Quadratic Gaussian (LQG) control is then presented as the linear combination of the two, with further details related to the separation principle. All these approaches will also be directly linked to different models used in perceptual and motor neuroscience, extensively reviewed in Chapter 2.

7.2.1 Linear Quadratic Estimator (LQE) or Kalman(-Bucy) filter

One of the most popular algorithms for estimation problems is the Kalman filter (KF) (Kalman, 1960b), or Kalman-Bucy filter for continuous systems (Kalman and Bucy, 1961), with applications spanning the most diverse fields, see for instance Chen (2003) and references therein. The Kalman(-Bucy) filter, also known as Linear Quadratic Estimator (LQE), is especially popular since 1) it's optimal for the estimation of linear systems, 2) it provides an advantageous iterative algorithmic implementation and 3) it is defined without specific assumptions on the stationarity of a system (Sorenson, 1970; Jazwinski, 1970; Chen, 2003). In this treatment we will use the continuous formulation and therefore focus on the Kalman-Bucy implementation. To define the filter, we initially describe a (linear) continuous dynamical systems in a state-space form, representing the system whose state(s) the filter is trying to estimate:

$$\begin{aligned} d\boldsymbol{x} =& A\boldsymbol{x}\, dt + d\boldsymbol{w} \\ \boldsymbol{y} =& C\boldsymbol{x} + d\boldsymbol{z} \end{aligned} \tag{7.1}$$

Here the bold characters indicate vectors. The first equation describes linear dynamics for a vector of processes $\boldsymbol{x}$. In the second equation, $\boldsymbol{y}$ is a vector of (noisy) measurements, or observations, of processes $\boldsymbol{x}$. Vectors $\boldsymbol{w}, \boldsymbol{z}$ are Wiener processes for state and observation equations respectively, with $d\boldsymbol{w} = \mathcal{N}(0, \Sigma_w), d\boldsymbol{z} = \mathcal{N}(0, \Sigma_z)$ thus defined as zero-mean white Gaussian random variables with covariance matrices Σ_w, Σ_z, represented according to Ito's definition of white noise relying on strictly zero-autocorrelation functions for $\boldsymbol{w}, \boldsymbol{z}$ (Jazwinski, 1970; Chen, 2003), see previous chapter for discussions on Ito/Stratonovich formulations. Matrices are represented using capital letters, C is the measurement/observation matrix mapping processes $\boldsymbol{x}$ to observations $\boldsymbol{y}$ and A is the state transition matrix characterising the dynamic behaviour of $\boldsymbol{x}$.

The Kalman-Bucy filter is well known in the literature to be the optimal estimator for linear systems with quadratic cost functions and Gaussian white random variables (Kalman, 1960b; Jazwinski, 1970; Chen, 2003; Åström and Murray, 2010).

It is also known to be a minimum variance estimator (Jazwinski, 1970; Chen, 2003), minimising the mean square error (MSE), or variance of the error, given by:

$$J = E[(\boldsymbol{x} - \hat{\boldsymbol{x}})^T(\boldsymbol{x} - \hat{\boldsymbol{x}})] \tag{7.2}$$

where $\hat{\boldsymbol{x}}$ is a vector containing the estimates of states $\boldsymbol{x}$. The Kalman-Bucy filter is usually presented as (Jazwinski, 1970; Stengel, 1994; Chen, 2003):

$$\begin{aligned}\dot{\hat{\boldsymbol{x}}} =& A\hat{\boldsymbol{x}} + K(\boldsymbol{y} - C\hat{\boldsymbol{x}})\\ K =& PC^T(\Sigma_z)^{-1}\\ \dot{P} =& \Sigma_w + AP + PA^T - K(\Sigma_z)K^T\end{aligned} \tag{7.3}$$

with the estimates $\hat{\boldsymbol{x}}$ given by the solution to the first equation, a sum of the current best estimate multiplied by the known transition matrix, $A\hat{\boldsymbol{x}}$ and prediction errors, or innovations, $(y - C\hat{\boldsymbol{x}})$. These prediction errors are multiplied by the so-called Kalman gain (matrix) K, described in the second equation, which represents the optimal trade-off between previous estimates and new information gathered from new observations. To calculate the Kalman gain matrix it is necessary to estimate P, the a posteriori error covariance matrix, expressing the accuracy of the state estimate in the first equation. The trace of P (i.e., the sum of the components on the main diagonal) gives the sum of the independent components of the covariance matrix (i.e., the sum of the variances of the single independent errors) equal to the MSE in equation (7.2). Kalman(-Bucy) equations thus minimise the MSE expressed in equation (7.2), with small values in the error covariance matrix implying high accuracy of the state estimation process. The process of estimation based on Kalman(-Bucy) filters can also easily be seen in terms of Bayesian inference processes, where a multivariate Gaussian distribution $\boldsymbol{x}$ is estimated using the above equations to create another multivariate Gaussian distribution with means $\hat{\boldsymbol{x}}$ and covariance matrix P (Meinhold and Singpurwalla, 1983; Chen, 2003). The general structure provided in these models gives rise to a series of hypotheses regarding Bayesian interpretations of perception in predictive coding (Rao and Ballard, 1999; Rao, 1999) and is then generalised to control and behaviour, for instance, in active inference and the free energy principle (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston, 2008a; Friston, 2010b), see Chapter 2.

Linear Quadratic Regulator (LQR) The Linear Quadratic Regulator (LQR) is the most basic example of closed-loop control found in the optimal control literature (Anderson and Moore, 1990; Stengel, 1994; Åström and Murray, 2010). In its simplest form, it is defined for deterministic linear systems with quadratic cost functions, for which the optimal control law turns out to be a simple negative feedback mechanism

(Anderson and Moore, 1990) in the case of infinite-horizon control. Given a system:

$$d\boldsymbol{x} = A\boldsymbol{x}\, dt + B\boldsymbol{a}\, dt \tag{7.4}$$

representing a linear, noiseless, environment, we define $\boldsymbol{x}$ as a vector of measured variables to be controlled. In this case the vector of processes to control are assumed to be directly observable (no measurement noise, $d\boldsymbol{z}$ in the LQE formulation) so a formulation in terms of Ito's calculus (using the differential form $d\boldsymbol{x}$) is not strictly necessary but maintained for consistency with the Kalman-Bucy filter definition provided previously. The vector $\boldsymbol{a}$ defines the actions that can be applied to $\boldsymbol{x}$. In more traditional formulations, this vector is usually represented by $\boldsymbol{u}$, but in this work we want to highlight a difference that will become crucial later. I will define $\boldsymbol{u}$ as a vector of inputs from a more general state-space perspective, i.e., variables that affect the state of a system but are not states themselves. On the other hand, $\boldsymbol{a}$ represents a subset of inputs $\boldsymbol{u}$, i.e., the (motor) actions of an agent, while $\boldsymbol{I}$ is used (often implicitly) for external forces generated by the world. In (optimal) control theory, these two definitions are often used interchangeably, assuming that all inputs are motor actions and no external forces can affect the state of a system. In our case however, we will also consider exogenous inputs/forces from the environment ($\boldsymbol{I}$) which are not represented by an agent's actions but can still affect its state (Sontag, 2003; Åström and Murray, 2010). A is the state transition matrix, as in the case of the Kalman-Bucy filter, while B is a matrix mapping actions $\boldsymbol{a}$ to outputs $\boldsymbol{x}$.

In LQR, the goal is to stabilise (control or regulate) variables $\boldsymbol{x}$ around target values $\bar{\boldsymbol{x}}$ (for simplicity here we will assume $\bar{\boldsymbol{x}} = \boldsymbol{0}$) by controlling their behaviour through actions $\boldsymbol{a}$. Such actions are determined via the optimisation of a function that accumulates costs over time called cost-to-go or value function:

$$J = \int_0^\infty \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \frac{1}{2}\boldsymbol{a}^T R\boldsymbol{a}\, dt \tag{7.5}$$

which represents the infinite horizon simplification of the problem (i.e., the upper limit of the integral is infinity). The instantaneous version, simply called cost function or cost rate, is defined for LQR as:

$$c(\boldsymbol{x}, \boldsymbol{a}) = \frac{1}{2}\boldsymbol{x}^T Q\boldsymbol{x} + \frac{1}{2}\boldsymbol{a}^T R\boldsymbol{a} \tag{7.6}$$

with $Q \geq 0$ and $R > 0$ as arbitrary matrices representing the relative balance between the minimisation of the distance from the target and costs for control respectively, e.g., physical costs such as energy. The optimal action vector $\boldsymbol{a}$ can then be shown to be equal to (Anderson and Moore, 1990; Stengel, 1994; Todorov, 2006;

Kappen, 2011):

$$\begin{aligned}\boldsymbol{a} =& -L\boldsymbol{x} \\ L =& R^{-1}B^TV \\ -\dot{V} =& Q + A^TV + VA - L^TR^{-1}L\end{aligned} \tag{7.7}$$

where $\boldsymbol{a}$ essentially implements a negative feedback mechanism on $\boldsymbol{x}$ and L is **feedback gain** matrix, described in the second equation. The third equation defines the matrix V, which is the Hessian of the cost-to-go function in equation (7.5). The use of LQR controllers is central in several models of action and behaviour in the form of negative feedback/delta-rule mechanisms (Widrow and Hoff, 1960; Rescorla, Wagner, et al., 1972), see also (Todorov and Jordan, 2002; Li and Todorov, 2004; Todorov, 2005; Stevenson et al., 2009; Nagengast, Braun, and Wolpert, 2010; Yeo, Franklin, and Wolpert, 2016) for work explicitly mentioning LQR and its generalisation LQG (defined in the next section).

One of the limitations of LQR controllers, however, lies in the fact that they do not explicitly deal with state/observation uncertainty or noise, i.e., the original formulation is defined for deterministic systems. On the other hand, in real-world engineering applications as well as in biological systems, it is in fact more common to think of systems with limited access to information from the environment and actions $\boldsymbol{a}$ thus applied to a set of hidden states $\boldsymbol{x}$ with only (noisy) measurements/observations $\boldsymbol{y}$ available. In the control theory literature, Linear Quadratic Gaussian (LQG) control (Anderson and Moore, 1990; Stengel, 1994; Åström and Murray, 2010) gracefully combines estimation or inference and control for linear systems. Under a set of assumptions, LQG controllers can be seen as modular: estimation and control components can be built and optimised as (nearly) independent processes under the "separation principle".

7.2.2 Linear Quadratic Gaussian (LQG) control

Following the separation principle, the LQG controller produces optimal estimation and optimal control for linear systems, sequentially combining two separate sub-systems, a Kalman-Bucy filter and LQR, in an optimal (i.e., minimum-variance) way (Anderson and Moore, 1990; Stengel, 1994; Kappen, 2011). The Kalman-Bucy filter provides the optimal state-estimate of an observation and the LQR controller uses such estimate (i.e., the mean) to implement the optimal deterministic controller: LQG control makes use of the estimated mean and feeds it into an LQR controller.

A general linear system to be regulated in the presence of noise on the observed state is described by:

$$\begin{aligned}d\boldsymbol{x} =& A\boldsymbol{x}\,dt + B\boldsymbol{a}\,dt + d\boldsymbol{w} \\ \boldsymbol{y} =& C\boldsymbol{x} + d\boldsymbol{z}\end{aligned} \tag{7.8}$$

where all the variables and parameters are the same as previously defined for Kalman-Bucy filters and LQR. In this case, the cost rate is slightly different as we are dealing with a stochastic system with white noise on both dynamics and observations. The standard approach of optimal control is thus extended to include stochastic variables and what is minimised is the *expected* cost-to-go (Stengel, 1994; Todorov, 2006), with cost rate defined as:

$$c(\boldsymbol{x}, \boldsymbol{a}) = \mathbb{E}\left[\frac{1}{2}\boldsymbol{x}^T Q \boldsymbol{x} + \frac{1}{2}\boldsymbol{a}^T R \boldsymbol{a}\right] \tag{7.9}$$

Using the separation principle, it can then be shown that minimising the expected value of the cost-to-go is equivalent to minimising the cost-to-go for the expected state (Kappen, 2011)

$$c(\boldsymbol{x}, \boldsymbol{a}) = c(\hat{\boldsymbol{x}}, \boldsymbol{a}) = \frac{1}{2}\hat{\boldsymbol{x}}^T Q \hat{\boldsymbol{x}} + \frac{1}{2}\boldsymbol{a}^T R \boldsymbol{a} \tag{7.10}$$

where we replaced states $\boldsymbol{x}$ with their estimates $\hat{\boldsymbol{x}}$, meaning that the optimal control can be computed using only the state estimate (i.e., the mean) $\hat{\boldsymbol{x}}$. The problem of estimation and control in LQG terms is then implemented by the following system combining Kalman-Bucy filter and LQR equations:

$$\begin{aligned}
\dot{\hat{\boldsymbol{x}}} =& A\hat{\boldsymbol{x}} + B\boldsymbol{a} + K(y - C\hat{\boldsymbol{x}}) \\
\boldsymbol{a} =& - L\hat{\boldsymbol{x}} \\
K =& PC^T(\Sigma_z)^{-1} \\
L =& R^{-1}B^T V \\
\dot{P} =& \Sigma_w + AP + PA^T - K(\Sigma_z)K^T \\
-\dot{V} =& Q + A^T V + VA - L^T RL
\end{aligned} \tag{7.11}$$

The ability to easily evaluate the cost rate in equation (7.10) is the core feature of LQG architectures. In more intuitive terms, after starting from the probabilistic control problem in equation (7.9), which accounts for the limited knowledge of a system, the problem becomes essentially deterministic in equation (7.10), where we simply replaced the state of a system with its estimate (see Stengel (1994) and Kappen (2011) for the mathematical proof). The state estimate is obtained via the use of a Kalman-Bucy filter, which is optimal and only conditionally dependent on the vector of motor actions $\boldsymbol{a}$ that contributed to the generation of observations $\boldsymbol{y}$ at the time of the estimate, i.e., neither the past nor the future of $\boldsymbol{a}$ are important. In this sense, the processes of estimation and control, or perception and action, exchange some basic information (estimates $\hat{\boldsymbol{x}}$ to the controller and actions $\boldsymbol{a}$ to the estimator) but are otherwise working independently. Moreover, the regulation problem is probabilistic only for the estimator, while a simple LQR can be used for control, as if the problem was deterministic and all information about a system was available to begin with.

The restrictive nature of the requirements for the separation principle and the related LQG architecture have been widely debated in the control theory literature (Stengel, 1994; Åström and Murray, 2010). Here we are especially interested and will discuss their possible meaning for the study of cognition, arguing that cognitivist accounts on the modular nature of the mind can be seen in analogy with the separation advocated by this principle. Countless studies in neuroscience and psychology make use of this assumption, focusing on either aspects of perception (Fodor, 1983; Marr, 1982; Knill and Richards, 1996; Rao, 1999) or action (Wolpert and Ghahramani, 2000; Todorov and Jordan, 2002; Todorov, 2004; Franklin and Wolpert, 2011), as if they could be modelled independently. However, living organisms are highly nonlinear and the environments they operate within are also themselves nonlinear, hence they cannot be fully understood with systems of linear equations (see assumption 1. in section 7.2). There is no evidence for noise in physical systems to be described by *white* Gaussian random variables (assumption 2.), on the other hand, it is often claimed that the assumption for noise to be white is not realistic (Van Kampen, 1981; Fox, 1987; Friston, 2008a). Whether biological system could effectively keep an updated estimate of the uncertainty in environmental variables (represented by covariance matrices in control theory) is also unclear (assumption 3.), especially considering non-stationary environments (Åström and Murray, 2010). It is then non-trivial to describe biological phenomena with quadratic cost functions (assumption 4.) (Körding and Wolpert, 2004; Franklin and Wolpert, 2011). Lastly, the separation principle suggests that perceptual systems must have access to an accurate copy of outgoing motor commands and all other possible inputs from the environment (assumption 5.).

As pointed out above and as shown in Fig. 7.1, under this scheme, the estimator (i.e., the Kalman-Bucy filter) and the controller (i.e., LQR) exchange information in two ways. The estimator relays accurate estimates $\hat{\boldsymbol{x}}$ of latent variables in the world to the controller, which in turn sends a copy of the motor command $\boldsymbol{a}$ back to the estimator. This copy of the motor command is crucial to allow the estimator, used as a metaphor for sensory systems, to discount sensory consequences of motor actions. In the absence of this information, estimates of world variables quickly become imprecise and subsequently controls become unstable (Friston, 2011). The notion of a copy of motor signals is consistent with the classical idea of efference copy in neuroscience (Holst and Mittelstaedt, 1950; Crapse and Sommer, 2008; Schwartz, 2016; Straka, Simmers, and Chagnaud, 2018). Efference copy is thought to represent a copy of signals from low-level motor areas in the brain that is sent to processing areas in order to disambiguate movements performed by an agent from environmental stimuli, although its definition is often vague and mixed with the idea of corollary discharge (Crapse and Sommer, 2008; Schwartz, 2016; Straka, Simmers, and Chagnaud, 2018). In the most prominent examples of LQG-based architectures in the cognitive sciences, efference copy is necessary for appropriate estimations of hidden variables in the world (Kawato, 1999; Wolpert and Ghahramani, 2000; Todorov, 2004; Wolpert, Diedrichsen, and Flanagan, 2011). Its exact definition and plausibility in neural system have however been comprehensively challenged (Feldman, 2009; Friston, 2011; Adams, Shipp, and Friston, 2013; Feldman et al., 2015; Feldman, 2016). In particular, some authors claim that the neurophysiological evidence supporting its existence is conflicting (Feldman, 2009; Adams, Shipp, and Friston, 2013; Feldman et al., 2015; Feldman, 2016), proposing then alternative models that eschew this idea entirely (Friston, 2011; Adams, Shipp, and Friston, 2013). Furthermore, this still doesn't speak to the presence of external forces affecting the observations of an agent and that cannot be known for different reasons. On the other hand, the presence of efficient mechanisms to counteract the effects of unknown stimuli is a defining feature of biological systems (Sontag, 2003) and remains unexplained by LQG architectures.

While the separation principle may not strictly hold, one could argue for a weaker version of separability/modularity and claim, nonetheless, its validity at least as a general driving theory for the study of cognitive systems. Agents could be cast as "approximately" or "partially" separable, following a less strict definition of separation (see Whittle (1981) for instance), with such notions still useful metaphors to understand perception and action as separate modules even without optimality as defined by the separation principle. This is especially true for the first four assumptions of the separation principle that we listed. For example, it may be possible to approximate nonlinear descriptions of a system with linearisations around relevant points/equilibria of a system (assumption 1.) (Åström and Murray, 2010) or to describe coloured noise as a high order autoregressive process expressed in terms of white noise (assumption 2.) (Chui and Chen, 2017). One may also think of estimating covariance matrices that although not optimal, closely resemble the uncertainty of a process (assumption 3.) (Rao, 1999). It may also be that quadratic cost functions provide good approximations in some instances (assumption 4.) (Simon, 1956; Körding and Wolpert, 2004). On the other hand, we argue that a weaker notion of "approximate" separability is not well defined for the last of the requirements of the separation principle (assumption 5.). This will thus be the main focus of our argument for the cognitive sciences.

The assumption that an estimator needs to have information about its own motor actions and other forces from the environment is often not considered a problem in control theory and robotics: a copy of motor signals can easily be retrieved and sent back to the estimator/forward model (Kawato, 1999) and external stimuli can be discounted by the experimenter. In biology and neuroscience, however, while the presence of information flowing from motor to sensory areas has been established for decades in the form of efference copy/corollary discharge (Cullen, 2004; Crapse and Sommer, 2008; Schwartz, 2016; Straka, Simmers, and Chagnaud, 2018), recent discussions on the information contents of such mechanisms (Feldman, 2009; Friston, 2011; Adams, Shipp, and Friston, 2013; Feldman, 2016) lead us to carefully consider frameworks based on these ideas and their role in the cognitive and natural sciences. Furthermore, modern experimental set ups in biology and neuroscience where unexpected external inputs are essentially denied have been put into discussions by several authors (Ahissar and Assa, 2016; Krakauer et al., 2017; Busse et al., 2017; Najafi and Churchland, 2018), questioning their relevance for ethologically meaningful explanations of natural behaviour. An alternative modelling approach, disposing with the need for a copy of motor signals and access to knowledge of all external inputs, is proposed with active inference. In active inference these issues are bypassed using a more powerful forward, or generative model, and trivial sensorimotor mappings in the form of reflex arcs replacing complex inverse models/controllers (see Fig. 7.2), similar to threshold or referent control ideas (Feldman et al., 2015).

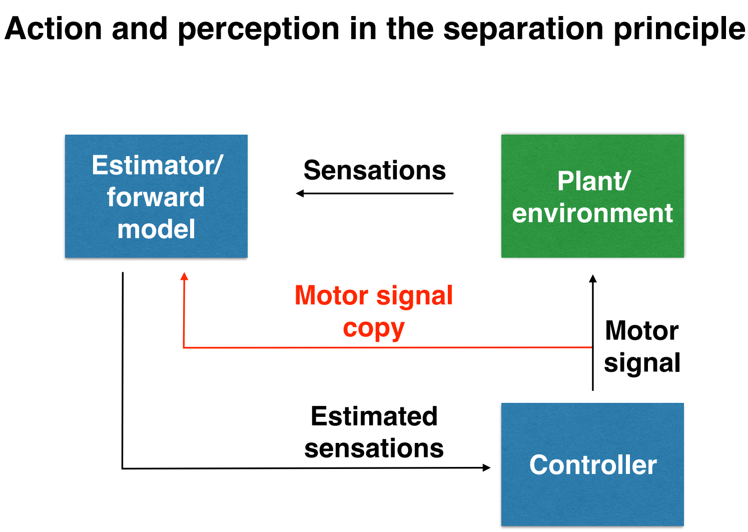


FIGURE 7.1: **A control architecture based on the separation principle.** An estimator, or forward model, infers the causes of incoming sensory input which are relayed to the controller. The controller calculates the optimal output motor signal based on these estimates and allows the system to act on the environment. In parallel, the controller sends also a copy of the command to the estimator, allowing the latter to take this command into account during the estimation of observed stimuli and discount the effects of internally generated actions.

7.3 Active inference and non-modular architectures

Active inference is a process theory proposed to explain brain functioning and other functions of living systems based on Bayesian inference and optimal control theory (Friston et al., 2010a; Buckley et al., 2017). In this section we establish its relations to the LQG architecture, starting by building an active inference version of the regulation of a linear multivariate system, and highlighting differences, limitations and possible extensions proposed for the control problem. As with LQG control, we will build an estimator of the hidden states $\boldsymbol{x}$. In this case however, we will give a variational account of the estimator in generalised coordinates of motion that generalises the MLE/MAP derivation of Kalman-Bucy filters (Jazwinski, 1970; Meinhold and Singpurwalla, 1983; Chen, 2003) using Variational Bayes with a Laplace approximation (Friston et al., 2007; Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston et al., 2010b; Buckley et al., 2017), see Chapter 3. We start by defining a generative model for an agent capturing the dynamics of the system to control and how these relate to observations and represented in a *generalised* state-space form[1] (Friston, 2008a; Buckley et al., 2017):

$$
\begin{aligned}
\boldsymbol{x}' &= \hat{A}\boldsymbol{x}' + \hat{B}\boldsymbol{v} + \boldsymbol{w} \\
\boldsymbol{y} &= \hat{C}\boldsymbol{x} + \boldsymbol{z}
\end{aligned}
\tag{7.12}
$$

where the hat over the matrices simply represents the fact that the matrices used in the generative model don't necessarily mirror their counterparts describing the world dynamics, as shown in our model later. The main difference with respect to LQG however, is that LQG explicitly mirrors (by construction in the linear case) the dynamics of the observed system, thus including knowledge of inputs $\boldsymbol{a}$, while in active inference this vector is not explicitly modelled by an agent, assuming that no copy of motor signals is available. It is in fact proposed that a deeper duality of estimation and control exists whereby, in the simplest case (i.e., a purely reflexive account), actions are just responses to the presence of prediction errors at the proprioceptive level, irrespectively of the cause of sensations, self-generated or external (Friston, 2011; Friston, Samothrakis, and Montague, 2012b; Brown et al., 2013; Adams, Shipp, and Friston, 2013). In recent accounts of more complex behaviour under active inference, action is cast as a problem of inference with (fictitious) control states $\boldsymbol{c}$ or rather time-dependent policies $\pi_{\boldsymbol{c}}$[2] that are inferred via the minimisation of expected free energy $\boldsymbol{y}$ (Friston, Samothrakis, and Montague, 2012b; Friston et al., 2015). Both these proposals support theories in motor neuroscience suggesting that knowledge of such self-produced controls (i.e., efference copy (Holst and Mittelstaedt, 1950)) is not available, and not necessary, for estimation in biological systems (Feldman, 2009; Friston, 2011; Adams, Shipp, and Friston, 2013; Feldman, 2016; Wiese, 2016). The vector $\boldsymbol{v}$ replacing $\boldsymbol{a}$ in the generative model encodes instead external or exogenous inputs in a state-space models sense. For the regulation of a system, these inputs represent priors or "desired" outcomes for an agent that could be derived from hierarchical implementations as outputs from layers above (Friston, 2008a).

[1] The notation for generalised coordinates introduced in Chapter 3, e.g., $\tilde{\boldsymbol{x}}$, is however dropped here for clarity. Moreover, for a fair comparison with LQG, the world dynamics of the model introduced later on are described using white noise, thus not requiring descriptions in terms of higher embedding orders. We will however maintain a Stratonovich definition of noisy terms, using for instance $\boldsymbol{w}$ rather than the differential form $d\boldsymbol{w}$.

[2] called $\boldsymbol{u}$ and $\pi_{\boldsymbol{u}}$ in (Friston et al., 2015)

This state-space model can then be written down in a probabilistic form, mapping the measurements equation to a likelihood $p(\boldsymbol{y}|\hat{\boldsymbol{x}})$ (no direct influence of inputs on observations), and the system's dynamics to a prior $p(\boldsymbol{x}, \boldsymbol{v})$ (Friston, 2008a), see also Chapter 3. The two probabilities densities are both Gaussian and can be written as:

$$
\begin{aligned}
p(\boldsymbol{y}|\boldsymbol{x}) &= \frac{1}{\sqrt{(2\pi)^m |\Sigma_z|}} \exp\left(-\frac{1}{2}(\boldsymbol{y} - \hat{C}\boldsymbol{x})^T \Sigma_z^{-1} (\boldsymbol{y} - \hat{C}\boldsymbol{x})\right) \\
p(\boldsymbol{x}, \boldsymbol{v}) &= \frac{1}{\sqrt{(2\pi)^n |\Sigma_w|}} \exp\left(-\frac{1}{2}(\boldsymbol{x'} - \hat{A}\boldsymbol{x'} - \hat{B}\boldsymbol{v})^T \Sigma_w^{-1} (\hat{\boldsymbol{x}}' - \hat{A}\boldsymbol{x'} - \hat{B}\boldsymbol{v})\right)
\end{aligned}
\tag{7.13}
$$

where m, n represent the number of elements of vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ respectively and $|\Sigma_z|, |\Sigma_w|$ are the determinants of the respective covariance matrices. With the general formulation of Laplace encoded variational free energy defined in Chapter 3 and here extended to the multivariate case:

$$
F \approx -\ln p(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{\theta}, \gamma)\Big|_{\boldsymbol{\vartheta}=\boldsymbol{\mu}} \tag{7.14}
$$

the free energy for a generic linear multivariate system becomes then:

$$
\begin{aligned}
F \approx &\frac{1}{2}\Bigg[\left(\boldsymbol{y} - \hat{C}\boldsymbol{\mu}_{\boldsymbol{x}}\right)^T \Pi_z \left(\boldsymbol{y} - \hat{C}\boldsymbol{\mu}_{\boldsymbol{x}}\right) + \left(\boldsymbol{\mu}'_{\boldsymbol{x}} - \hat{A}\hat{\boldsymbol{\mu}}_{\boldsymbol{x}} - \hat{B}\boldsymbol{\mu}_{\boldsymbol{v}}\right)^T \Pi_w \left(\boldsymbol{\mu}'_{\boldsymbol{x}} - \hat{A}\boldsymbol{\mu}_{\boldsymbol{x}} - \hat{B}\boldsymbol{\mu}_{\boldsymbol{v}}\right) \\
&- \ln\left(\Pi_z\right) - \ln\left(\Pi_w\right) + (m+n)\ln 2\pi\Bigg]
\end{aligned}
\tag{7.15}
$$

where we defined precision matrices Π_z, Π_w as the inverse of covariance matrices Σ_z, Σ_w. Variables m, n represent the dimensions of vectors $\boldsymbol{y}$ and $\boldsymbol{x}$ respectively. It is important to highlight that, in general, the covariance matrices used in the generative model can be different from the ones used to describe the environment or generative process. To simplify the already heavy notation we will however represent them in the same way. We also explicitly replaced $\boldsymbol{x}$ with their expectations $\boldsymbol{\mu}_{\boldsymbol{x}}$, since under the Laplace assumption this represents the best estimate of $\boldsymbol{x}$. Expectations $\boldsymbol{\mu}_{\boldsymbol{x}}$ play the same role of estimates $\hat{\boldsymbol{x}}$ in LQG, we simply decided to use a notation consistent with some of our previous work, see Buckley et al. (2017) and previous chapters.

The recognition dynamics, encoding perception and action in a system minimising free energy (Friston, Trujillo-Barreto, and Daunizeau, 2008; Friston et al., 2010b; Buckley et al., 2017) and equivalent to estimation and control functions respectively, are implemented in standard active inference formulations as a gradient descent scheme minimising the free energy with respect to the means $\hat{\boldsymbol{x}}$ for perception/estimation:

$$
\begin{aligned}
\dot{\boldsymbol{\mu}}_{\boldsymbol{x}} &= \boldsymbol{D}\boldsymbol{\mu}_{\boldsymbol{x}} + \hat{C}^T \Pi_z \left(\boldsymbol{y} - \hat{C}\boldsymbol{\mu}_{\boldsymbol{x}}\right) + \hat{A}^T \Pi_w \left(\boldsymbol{\mu}'_{\boldsymbol{x}} - \hat{A}\boldsymbol{\mu}_{\boldsymbol{x}} - \hat{B}\boldsymbol{\mu}_{\boldsymbol{v}}\right) \\
\dot{\boldsymbol{\mu}}'_{\boldsymbol{x}} &= \boldsymbol{D}\boldsymbol{\mu}'_{\boldsymbol{x}} - \Pi_w \left(\boldsymbol{\mu}'_{\boldsymbol{x}} - \hat{A}\boldsymbol{\mu}_{\boldsymbol{x}} - \hat{B}\boldsymbol{\mu}_{\boldsymbol{v}}\right)
\end{aligned}
\tag{7.16}
$$

and actions $\boldsymbol{a}$ for action/control, assuming only that actions have an effect on observations $\boldsymbol{y}$:

$$\dot{\boldsymbol{a}} = -\frac{\partial F}{\partial \boldsymbol{a}} = -\frac{\partial F}{\partial \boldsymbol{y}}\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{a}} = -\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{a}}^T \Pi_z \left(\boldsymbol{y} - \hat{C}\hat{\boldsymbol{x}}\right) \tag{7.17}$$

The estimation expressed in equation (7.16) prescribes a generalisation of Kalman-Bucy filters to trajectories with arbitrary embedding orders where random variables are treated as smooth stochastic processes, i.e., without the Markov assumption necessary in standard Kalman-Bucy filters (see above). In equation (7.16), we also include a term $\boldsymbol{D}\boldsymbol{\mu}_{\boldsymbol{x}}$, previously introduced in Chapter 3 as the mode (or mean for Gaussian variables) of the motion for the minimisation in generalised coordinates of motion (Friston, 2008a; Buckley et al., 2017). Action as expressed in (7.17) may appear similar to the traditional LQR/LQG form, but is fundamentally different since it depends on observations $\boldsymbol{y}$ rather than on estimates of hidden states $\boldsymbol{\mu}_{\boldsymbol{x}}$.

7.3.1 Action and perception are not separable

In active inference, perfect knowledge of the motor signals is not assumed to be necessary for combined models of control and estimation (Friston, 2011; Adams, Shipp, and Friston, 2013), and possibly not at all present in biological systems (Feldman et al., 2015; Feldman, 2016). The proposed alternative entails recasting motor control problems into perceptual or inference problems, considering that these two classes of problems can be solved by the same algorithms (Kalman, 1960c; Anderson and Moore, 1990; Attias, 2003; Todorov, 2008; Friston, 2011). According to active inference, perception and action are largely overlapping processes sharing most of their computation, with differences arising mainly at a physiological level (Adams, Shipp, and Friston, 2013). The problem of finding actions is essentially converted into an inference problem, solved by the same underlying predictive coding scheme implementing perceptual processes, see Fig. 7.2. In this architecture, proprioceptive sensations are also predicted by an agent, alongside exteroceptive and interoceptive ones. Explicit motor output is then produced by simple sensorimotor mappings implemented at the very periphery of a system, translating proprioceptive predictions into actions in the external world (Friston et al., 2010a). Conceptually, active inference disposes with the need of a copy of motor signals proposing a more general predictive coding scheme based on priors applied to proprioceptors, coupled to simple sensorimotor mappings translating proprioceptive predictions into actual actions (Friston, 2011; Adams, Shipp, and Friston, 2013).

More in detail, unlike LQG where the Kalman-Bucy filter is provided with a copy of motor controls $\boldsymbol{a}$, effectively discounted from the estimation of states $\hat{\boldsymbol{x}}$, in active inference these signals are not provided and the problem of control is almost entirely replaced by the generalised Kalman-Bucy filter expressed in equation (7.16). This new "filter" provides estimates $\boldsymbol{\mu}_{\boldsymbol{x}}$ biased towards the desires of an agent (Friston et al., 2010a), rather than an objective account of observations $\boldsymbol{y}$. Action is then

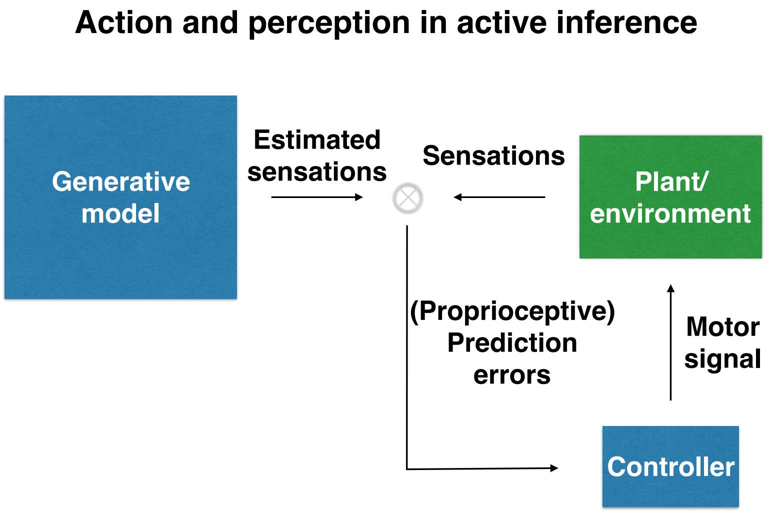


FIGURE 7.2: **A control architecture based on active inference.** Active inference converts the complex problem of optimal control into a more viable problem of inference, solved by a more general generative model (Friston, 2011). The forward, or generative, model produces estimates of the sensory input. The mismatch between these estimates and real sensory data (here represented as only proprioceptive but more in general including also exteroceptive and interoceptive) generates prediction errors that are used to update the generative model itself and thus infer the causes of sensory data when they are minimised. Propriocetive prediction errors are also explicitly minimised via simple reflex mechanisms implemented at the level of peripheral "controllers". These controllers receive information in an intrinsic frame of reference, proprioceptive signals within an agent, and translate them into controls in an extrinsic one, motor actions in the world, using hardwired sensorimotor mappings (Friston et al., 2010a).

determined entirely through its effects on observations $\partial \boldsymbol{y}/\partial \boldsymbol{a}$ instead of requiring precise knowledge on how control affects estimates of hidden states $\boldsymbol{\mu_x}$, thought to be a much harder problem to solve (Friston, 2011). In active inference, the more standard inverse problem used to formulate motor control problems in (Kawato, 1999; Wolpert and Ghahramani, 2000) as mappings is decomposed into two parts. The first one, mapping desires/reference values in extrinsic (i.e., world) coordinates to an internal frame of reference (i.e., proprioception) is solved by the same generative model already in charge of predicting and estimating hidden variables from observations of the real world. The second one, providing mappings between proprioceptive predictions and actions in the real world is thought to be implemented at the spinal level via simple reflex-like mechanisms (Friston, 2011; Adams, Shipp, and Friston, 2013), for simplicity assumed to be endowed by evolutionary processes rather than learnt (Friston et al., 2010a). In active inference, action is also expressed for systems in generalised coordinates of motion considering how, in the more general case, an agent's actions can affect a trajectory rather than a single order of motion, a feature that introduces for instance new interpretations of classical regulation strategies such as PID control (see Chapter 6).

In active inference, the more traditional, sequential and modular role of perception and action advocated by the separation principle has already been questioned, suggesting that these processes are deeply intertwined (Friston, 2011; Pickering and Clark, 2014; Engel, Friston, and Kragic, 2016; Wiese, 2016; Pezzulo et al., 2017). Essentially, this alternative architecture is based on a fundamentally entangled action-perception loop where the two processes cannot really be investigated independently, since action is enacted by systematic perceptual misrepresentations (Wiese, 2016) and classical accounts of perceptual functions only exist as a special, non-ecologically plausible instance of agent-environment systems (Friston, Thornton, and Clark, 2012). In support of this idea, and following our own argument on the parallelism between Fodor's idea of modularity and the separation principle of control theory, we claim that active inference does not formally meet the requirements for the separation of estimation and control (perception and action). Furthermore, it is engaging in an explicitly non-modular architecture of cognitive processing. As previously suggested, while the first four requirements of the separation principle may be prone to arguments regarding the existence of separation in at least some approximate sense (e.g., an approximately linear model, noise that is approximately Gaussian, etc.), the presence of a copy of motor signals generates a strong dichotomy, with no room for approximations. In this very specific case, the separation of estimation and control is only possible in presence of this copy and no approximate separation can exist when this is missing, since Kalman-Bucy are not well defined without this information (Kalman, 1960c; Chen, 2003). Without a copy of motor signals, the architecture described by the separation principle is unavoidably broken and thus, according to our initial claim regarding the cognitive sciences, Fodor's modularity cannot be implemented. More in general, the presence of any external force $\boldsymbol{I}$ is bound to generate the same effects: if inputs $\boldsymbol{I}$ are not represented within the Kalman-Bucy filter, the separation principle cannot be defined since *all* inputs need to be provided. Active inference models, by explicitly eschewing the idea that a copy of motor signals is sent to estimators (Friston, 2011), embrace a fully non-modular perspective with action and perception intimately entangled only "separated" in name for consistency with traditional labels used in cognitive science.

7.4 The model

The double integrator is a canonical example used to describe control theoretical methods and is one of the simplest and most fundamental problems in optimal control, modelling single degree-of-freedom motion of different physical systems (Rao and Bernstein, 2001; Åström and Murray, 2010). In the simplest case presented here, this could be thought of as, essentially, a block on frictionless surface. The standard double integrator is usually described as a deterministic system. The control policy is thus usually defined using a feedback law applied directly to the *known* dynamics, as the full state of the system is measured with no uncertainty (Rao and Bernstein, 2001). For the purposes of this work, where uncertainty and noise are crucial components, we will introduce process and measurement noise into the system, making the estimation of hidden states necessary and thus fully compare LQG and active inference in one of the simplest possible examples in the control theory literature[3]. The double integrator is described by the following state-space model:

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{a} + \boldsymbol{w} \\ \boldsymbol{y} &= C\boldsymbol{x} + \boldsymbol{z} \end{aligned} \tag{7.18}$$

where observations, hidden states, controls and noise/fluctuations are defined as vectors:

$$\boldsymbol{y} = \begin{bmatrix} y \\ y' \end{bmatrix} \quad \boldsymbol{x} = \begin{bmatrix} x \\ x' \end{bmatrix} \quad \boldsymbol{a} = \begin{bmatrix} 0 \\ a \end{bmatrix} \quad \boldsymbol{z} = \begin{bmatrix} z \\ z' \end{bmatrix} \quad \boldsymbol{w} = \begin{bmatrix} w \\ w' \end{bmatrix}$$

and matrices A, B, C as:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.19}$$

while covariance matrices Σ_z, Σ_w as:

$$\Sigma_z = \begin{bmatrix} \exp(0) & 0 \\ 0 & \exp(0) \end{bmatrix} \quad \Sigma_w = \begin{bmatrix} 0 & 0 \\ 0 & \exp(-1) \end{bmatrix} \tag{7.20}$$

which effectively reduces the system to

$$\begin{aligned} \dot{x} &= x' & y &= x + z \\ \dot{x}' &= x'' = a + w & \dot{y} = y' &= x' + z' \end{aligned} \tag{7.21}$$

[3]The code is available at https://github.com/mbaltieri/doubleIntegrator.

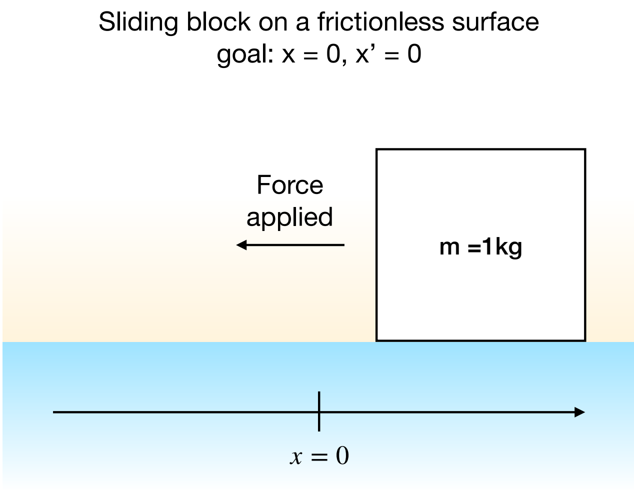


FIGURE 7.3: **The generative process, a double integrator.** The double integrator problem, corresponding to a block of mass=1kg placed on a surface with no friction. The goal is to have the block moving to position $x = 0$ and then stop $x' = 0$.

7.4.1 The LQG solution to the double integrator

With LQG, the equations for estimation and control in the double integrator following the Kalman-Bucy filter + LQR structure are (see equation (7.11)):

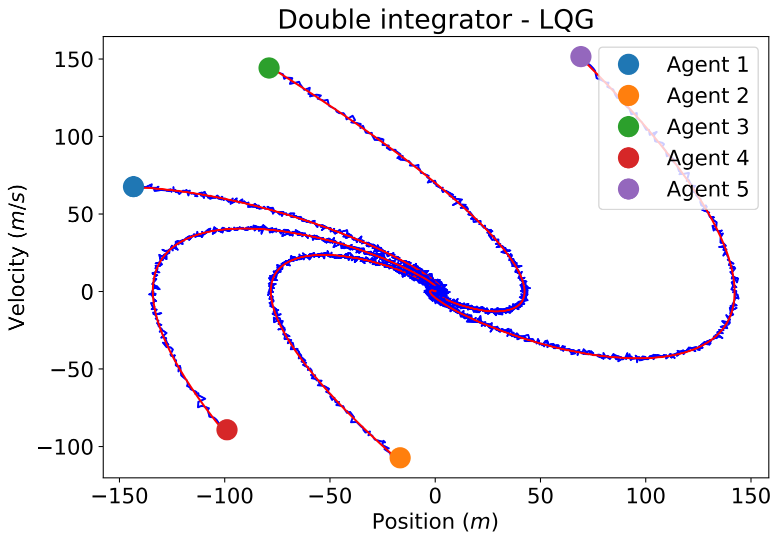
$$\begin{aligned}
\dot{\hat{x}} =& \hat{x}' + k_1(y - \hat{x}) + k_2(y' - \hat{x}') \\
\dot{\hat{x}}' =& a + k_3(y - \hat{x}) + k_4(y' - \hat{x}') \\
a =& -l_1(\hat{x} - r) - l_2(\hat{x}' - r') \\
K =& PC^T(\Sigma_z)^{-1} \\
L =& R^{-1}B^TV \\
\dot{P} =& \Sigma_w + AP + PA^T - K(\Sigma_z)K^T \\
-\dot{V} =& Q + A^TV + VA - L^TRL
\end{aligned} \tag{7.22}$$

where r is the target position, $r = 0$ for the double integrator. For the LQR component, we then specify matrices:

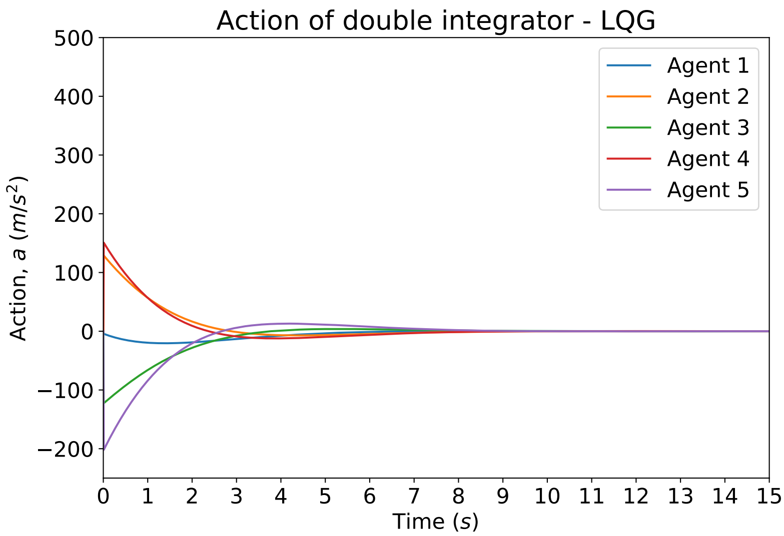
$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad R = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \tag{7.23}$$

not optimising their values since it is beyond the scope of this work. For further results and analysis see for instance (Rao and Bernstein, 2001).

As we can see in Fig. 7.4a, the block is effectively driven to the desired position $x = 0$ and velocity $x' = 0$ from a set of 5 randomly initialised conditions (zero-mean



(A)

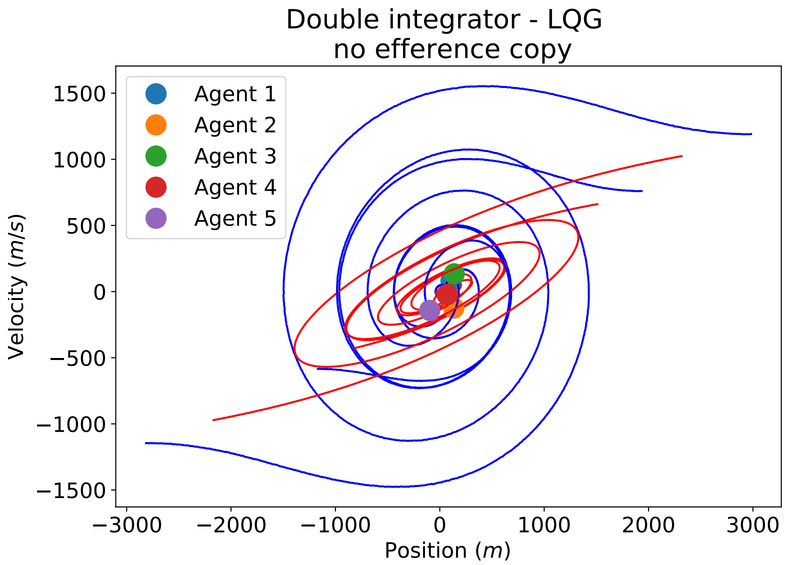


(B)

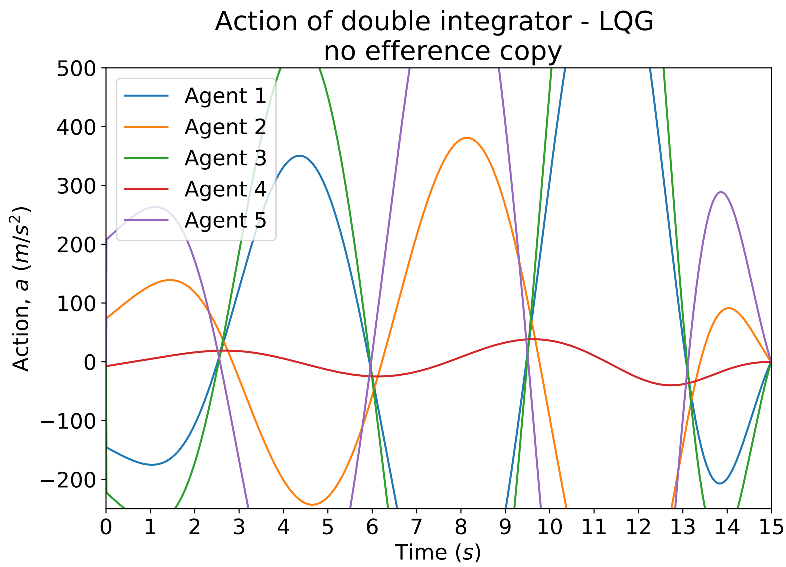
FIGURE 7.4: **The double integrator solved using LQG.** (a) Five LQG agents with different initial conditions showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents.

Gaussian distributed, sd=100). In Fig. 7.4b we then simply show the actions over time of the same 5 examples, all converging to zero since the agents effectively reach their desired target. The main feature of LQG, and from which active inference will depart, is the reliability of estimates of both position and velocity (the red line in the phase space). In LQG, accurate estimates are necessary to then enact the LQR component implementing a negative feedback mechanism based on estimates $\hat{\boldsymbol{x}}$ rather than true hidden states $\boldsymbol{x}$.

When knowledge of the motor signals $\boldsymbol{a}$ is removed from equation (7.11), estimates of the hidden properties of the world become inaccurate and unstable, as



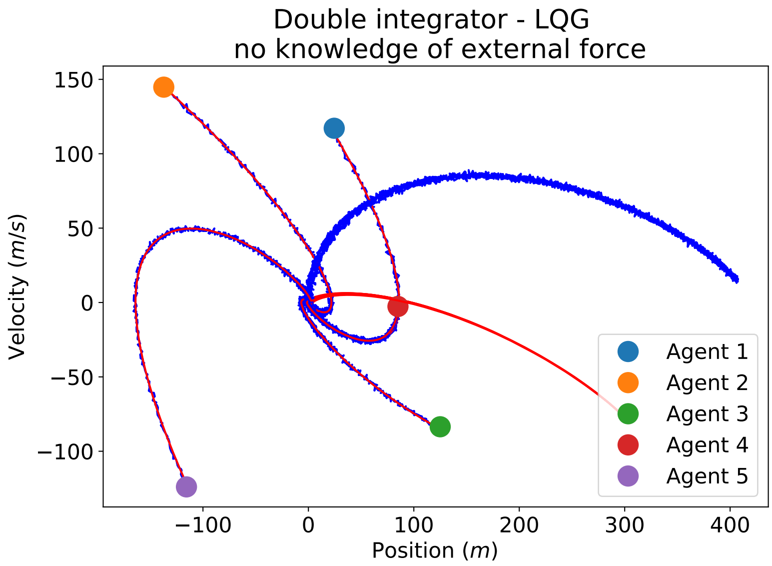
(A)



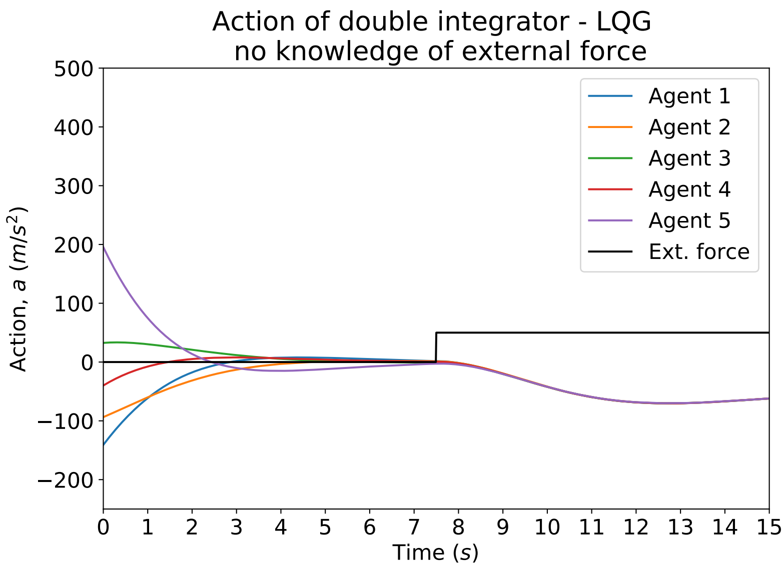
(B)

FIGURE 7.5: **The double integrator solved using LQG with no knowledge of motor signals.** (a) Five examples with different initial conditions showing in blue the observed trajectories of different blocks in the phase-space and in red the agent's estimates of the same trajectories. In this case we specifically removed motor signals $\boldsymbol{a}$ from equation (7.11), simulating a lack of information regarding these signals. (b) Actions taken by the five agents.

shown in Fig. 7.5 for the double integrator. In this example, rather than converging towards the desired state, our agents get away from it (Fig. 7.5a) since the new observations are too inaccurate given the lack of mechanisms to discount the effects of $\boldsymbol{a}$. In Fig. 7.5b we can then see that actions $\boldsymbol{a}$ begin to exponentially oscillate rather than converging to zero, as in Fig. 7.4. This is due to one of the assumptions for observability defined by Kalman in (Kalman, 1960c), explicitly requiring knowledge of *all* inputs and outputs of a system in order to determine its latent state(s).



(A)



(B)

FIGURE 7.6: **The double integrator solved using LQG with no knowledge of external forces.** (a) Five LQG agents with different initial conditions showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents (coloured lines) and external force applied to the agent (in black).

In Fig. 7.6 we introduced a new external force $\boldsymbol{I}$ not modelled by the agents, equivalent to a disturbance from the environment (black line in Fig. 7.6b). Fig. 7.6a then shows that the agents are incapable of regulating their position/velocity against this unknown input (blue lines), after an initial convergence towards the desired state, they in fact move away from the target when the force is introduced. Furthermore, these agents are incapable of correctly inferring their trajectories, providing inaccurate estimates of their sensed variables (red lines). In Fig. 7.6b we see that all of these agents attempt to counteract the effects of unexpected stimuli (they minimise their velocity after the force is introduced), however the lack of an appropriate mechanism to track their position correctly (i.e., integral action) pushes them away from the target.

7.4.2 The double integrator with active inference

To solve the same control problem, active inference relies on the generation of predictions of proprioceptive sensations (position, velocity and acceleration of the agent in this case), followed by the implementation of actions in the world via (trivial) reflex arcs. The proprioceptive modality is essentially treated as other inputs (vision, audition, etc.) and estimates/predictions are generated using the same generative model taking advantage of incoming proprioceptive sensations. This produces a considerably different control system, with state estimates and actions now created by the same model, making it hard to clearly separate processes of perception and action, see Fig. 7.2. The copy of motor control signals (cf. efference copy (Holst and Mittelstaedt, 1950)), necessary in standard LQG settings to meet the observability constraints of Kalman-Bucy filters (Anderson and Moore, 1990; Stengel, 1994) is not included in this formulation, as explained in section 7.3.1. Active inference postulates in fact that direct representations of the causes or actions $\boldsymbol{a}$ of self-generated sensations need not be discounted during the prediction of new incoming sensory inputs. This could be seen as a limitation of active inference accounts, but in general it speaks to the robustness of this approach in face of unknown inputs (i.e., motor actions produced by an agent or exogenous forces from the environment). It is also claimed that, in this framework, inputs/causes can be estimated using an appropriate generative model of the world dynamics (Friston, Trujillo-Barreto, and Daunizeau, 2008), a feature thought to be defining for biological systems (Sontag, 2003). Simple and effective approximations are also possible, for example with integral control, thought to be the most basic heuristic dealing with the problem of uncertain inputs even in biological systems down to the unicellular level (Sontag, 2003) and shown to be consistent with formulations of linear generative models in active inference (see Chapter 6). To derive the active inference solution to the same double integrator problem, we start by defining a generative model for the agent, i.e., the block:

$$\begin{aligned} \boldsymbol{x}' &= \hat{A}\boldsymbol{x} + \hat{B}\boldsymbol{v} + \boldsymbol{w} \\ \boldsymbol{y} &= \hat{C}\boldsymbol{x} + \boldsymbol{z} \end{aligned} \tag{7.24}$$

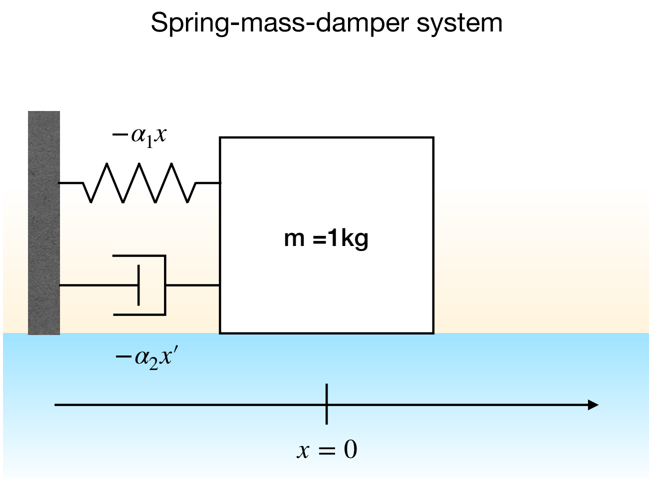


FIGURE 7.7: **The generative model for the double integrator.** The agent implements a generative model with priors representing an imaginary spring that pulls the block back to the origin ($x = 0$) and an imaginary damper that slows it down ($x' = 0$).

where observations, hidden states, inputs and noise/fluctuations are defined as vectors (for the notation, refer to Chapter 3):

$$\boldsymbol{y} = \begin{bmatrix} y \\ y' \\ y'' \end{bmatrix} \quad \boldsymbol{x} = \begin{bmatrix} x \\ x' \\ x'' \end{bmatrix} \quad \boldsymbol{v} = \begin{bmatrix} v \\ v' \\ v'' \end{bmatrix} \quad \boldsymbol{z} = \begin{bmatrix} z \\ z' \\ z'' \end{bmatrix} \quad \boldsymbol{w} = \begin{bmatrix} w \\ w' \\ w'' \end{bmatrix} \tag{7.25}$$

and matrices $\hat{A}, \hat{B}, \hat{C}$ as:

$$\hat{A} = \begin{bmatrix} 0 & 1 & 0 \\ -\alpha_1 & -\alpha_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \hat{B} = \begin{bmatrix} \exp(1) & 0 & 0 \\ 0 & \exp(1) & 0 \\ 0 & 0 & \exp(1) \end{bmatrix} \quad \hat{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{7.26}$$

with covariance matrices Σ_z, Σ_w defined as:

$$\Sigma_w = \begin{bmatrix} \exp(1) & 0 & 0 \\ 0 & \exp(1) & 0 \\ 0 & 0 & \exp(1) \end{bmatrix} \quad \Sigma_z = \begin{bmatrix} \exp(8) & 0 & 0 \\ 0 & \exp(8) & 0 \\ 0 & 0 & \exp(8) \end{bmatrix} \tag{7.27}$$

The description of this generative model already highlights profound differences with the LQG approach that will be discussed in detail in section 7.5. For now, it is important to notice that an extra observation is provided, the measured acceleration

y'', and that this generative model depicts the agent as a spring-mass-damper system. The agent implements beliefs of a world where it is pulled back to the desired state $x = 0$ and $x' = 0$ by an imaginary spring and slows down thanks to an imaginary piston-like damper, "designed" (in this case by us, but more in general one could imagine evolutionary processes for biological system (Friston et al., 2010a)) to favour normative behaviour.

Following equation (7.26), the variational free energy for our controller is then described by:

$$F \approx \frac{1}{2}\Big[\pi_z(y-\mu_x)^2 + \pi_{z'}(y'-\mu'_x)^2 + \pi_{z''}(y''-\mu''_x)^2 + \pi_{w'}(\mu''_x-\mu_v)^2 \tag{7.28}$$

$$- \ln(\pi_z\pi_{z'}\pi_{z''}\pi_{w'}) - 6\ln 2\pi\Big] \tag{7.29}$$

where precisions $\pi_z\pi_{z'}\pi_{z''}\pi_{w'}$ are taken from the diagonals of precision matrices Π_z, Π_w (inverse covariances matrices Σ_z, Σ_w defined in the generative model). After explicitly writing out the equations derived from the matrix formulation in (7.16), we get the following formulation of perceptual inference:

$$\begin{aligned}
\dot{\mu}_x &= \mu'_x + \pi_z(y-\mu_x) + \pi_w(\mu'_x + \alpha\mu_x - \beta\mu_v) \\
\dot{\mu}'_x &= \mu''_x + \pi_{z'}(y'-\mu'_x) + \pi_{w'}(\mu''_x + \alpha\mu'_x - \beta\mu'_v) \\
\dot{\mu}''_x &= \mu'''_x + \pi_{z''}(y''-\mu''_x) + \pi_{w''}(\mu''_x + \alpha\mu'_x - \beta\mu'_v)
\end{aligned} \tag{7.30}$$

and

$$\begin{aligned}
\dot{\mu}'_x &= -\pi_w(\mu'_x + \alpha\mu_x - \beta\mu_v) \\
\dot{\mu}''_x &= -\pi_{w'}(\mu''_x + \alpha\mu'_x - \beta\mu'_v)
\end{aligned} \tag{7.31}$$

showing another important difference derived from the lack of the Kalman gain K: if K is non-diagonal as in this case (one can simply verify this claim with standard functions solving continuous Riccati equations, as in the provided code), multiple orders of motion are present in the optimal filter problem in equation (7.11), but only one appears in equation (7.30) since the precision matrices are assumed to be diagonal in our formulation.

The action component is, however, the one most significantly different, starting from the assumption that direct knowledge of motor signals is not available and thus not modelled in the generative model (motor commands $\boldsymbol{a}$ are replaced by inputs $\boldsymbol{v}$ acting as priors). This entails a new approach to the understanding of the problem, with active inference suggesting that the only information needed comes from observations $\boldsymbol{y}$, see equation (7.17). On this account, action then reduces to

$$\dot{\boldsymbol{a}} = -\left(\frac{\partial \boldsymbol{y}'}{\partial a}\right)^T \Pi_z(\boldsymbol{y} - \hat{C}\boldsymbol{\mu}_x) \tag{7.32}$$

and with the assumption that

$$\frac{\partial \boldsymbol{y}}{\partial a} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$$

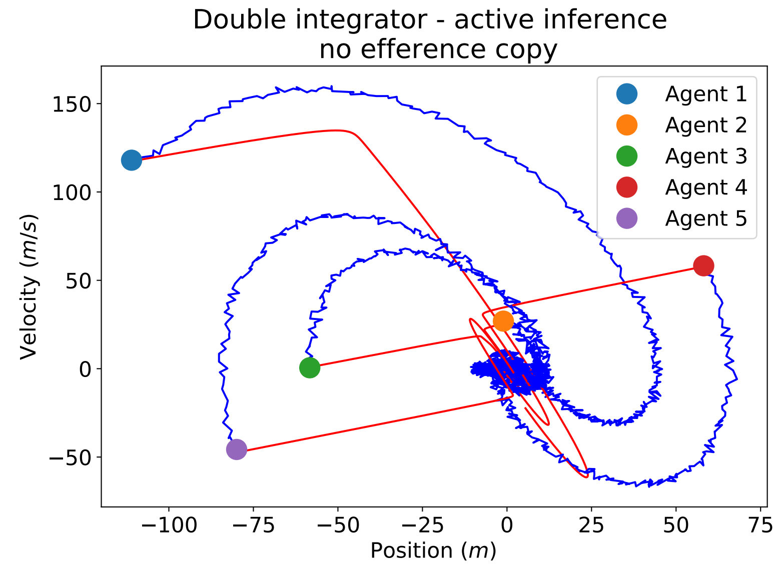
the explicit, scalar version of action becomes

$$\dot{a} = -\pi_z(y - \mu_x) - \pi_{z'}(y' - \mu'_x) - \pi_{z''}(y'' - \mu''_x) \tag{7.33}$$

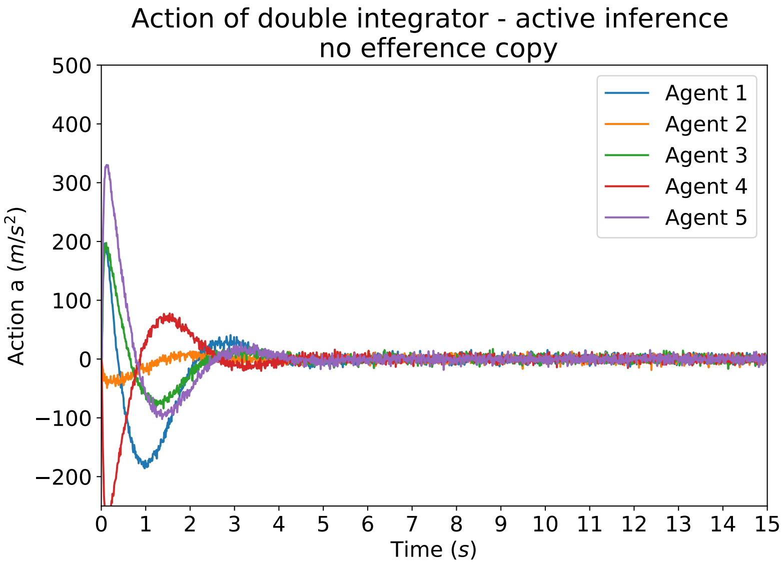
replacing the LQR component in equation (7.11). This type of control is equivalent to a PID controller (see Chapter 6) and as we will discuss later is the "optimal" linear solution when knowledge of inputs $\boldsymbol{a}$ or $\boldsymbol{I}$ is not available in the generative model. As in the case of filtering, the feedback gain L is missing in the active inference formulation, replaced by learning rates of the gradient descent or by clever approximations (see Chapter 4).

In Fig. 7.8 we can see an example implementation of the double integrator using active inference. Five agents are initialised at random position and velocity (zero-mean Gaussian distributed, sd=100) and Fig. 7.8a, and converge to the target solution where the output actions are essentially zero (excluding some noise), as expected Fig. 7.8b. The most striking feature is that estimates of both position and velocity of the block are very inaccurate but the agent nonetheless reaches the desired target in the phase space. These differences are given by the generative model implemented by the agent, encoding an imaginary spring-damper system that pulls it towards its desired state. By relaxing the stiffness of the spring and the friction represented by the damper we can see that the agents start circling around the target increasingly more and more, and are essentially slower at reaching their desired state, see Fig. 7.9a and Fig. 7.10a. However, their estimates of the latent variables, position and velocity, are more accurate. In active inference, action and perception can be seen as competing for the same resources (essentially an attention-like mechanism (Feldman and Friston, 2010; Brown et al., 2013)) within a single generative model, and better performances of one guarantee worse results for the other. Agents that are "too good" at inferring properties of their environment don't move (Brown et al., 2013), while agents that move too much might end up missing important details (Friston, Daunizeau, and Kiebel, 2009; Friston et al., 2010a): only a balanced mix of action and perception can allow for normative behaviour.

Fig. 7.11 shows then the robustness of this implementation when an external force $\boldsymbol{I}$ is introduced: by implementing integral control (Baltieri and Buckley, 2018a), active inference can in this case counteract the effects of unexpected inputs. The presence of integral action perfectly counteracts the effects of disturbances Fig. 7.11b (cf. Fig. 7.6b), and more importantly allows for the desired regulation of the agents' positions, Fig. 7.11a, which is impossible in LQG accounts assuming perfect knowledge of the world (cf. Fig. 7.6a).



(A)

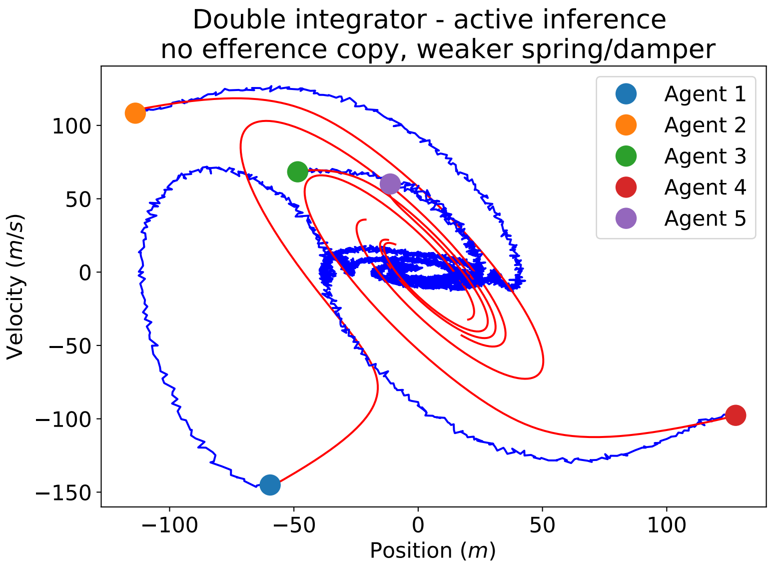


(B)

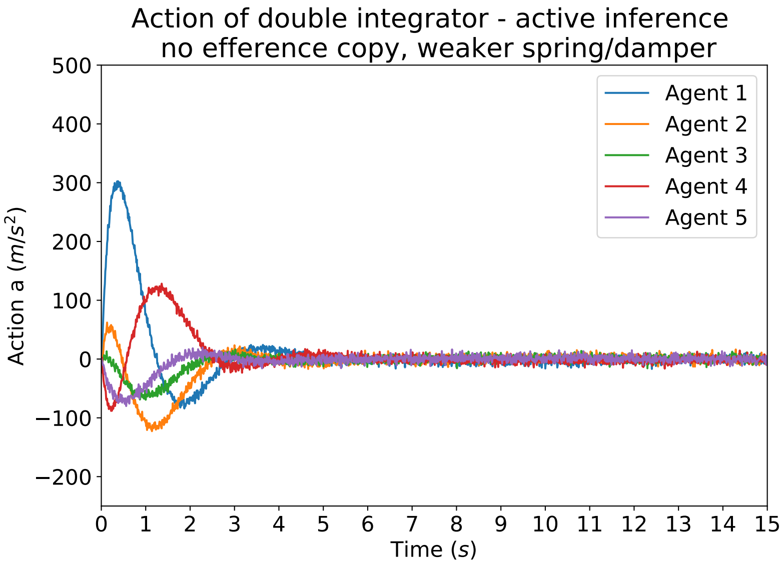
FIGURE 7.8: **The double integrator solved using active inference ($\alpha_1 = \exp(2), \alpha_2 = \exp(1)$).** (a) Five active inference agents with different initial conditions showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents.

7.5 Discussion

Active inference is an emerging theory in the cognitive sciences proposed to account for several phenomena of life and cognition, including action, perception, learning and other higher order functions (Friston, Kilner, and Harrison, 2006; Friston, 2010b; Friston et al., 2010a; Hohwy, 2013; Clark, 2015b). Its position over the long standing debate between computational and 4E (enactive, embodied, embedded, extended) theories of cognition remains however unclear (Clark, 2015a; Allen and



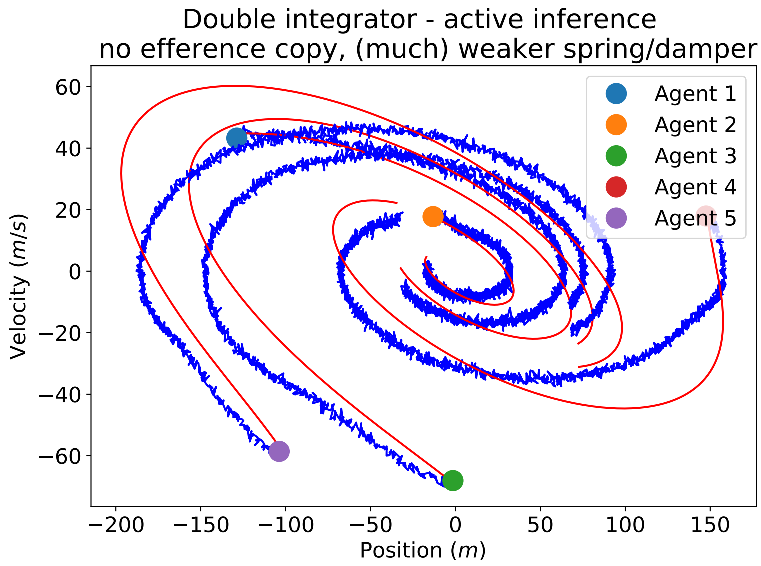
(A)



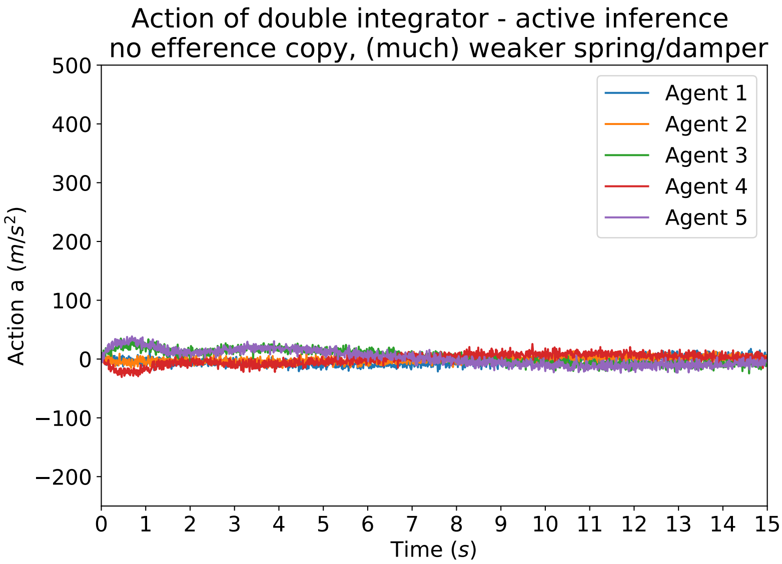
(B)

FIGURE 7.9: **The double integrator solved using active inference ($\alpha_1 = \exp(1), \alpha_2 = \exp(0.5)$).** (a) Five active inference agents with different initial conditions and different parameters of the imaginary mass-spring system implemented as a prior, $\alpha_1 = \exp(1), \alpha_2 = \exp(0.5)$. Showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents.

Friston, 2018; Bruineberg, Kiverstein, and Rietveld, 2018; Kirchhoff and Froese, 2017; Gładziejewski, 2016). In this work we focused on the idea of *modularity* explicitly introduced in the cognitive science debates by Fodor (1983). In particular, this idea is central to architectures based on the classical sandwich of cognitive science (Coltheart, 1999; Hurley, 2001; Barrett and Kurzban, 2006), where perception and action are seen as modules of a feedforward-only architecture that are explicitly separated by cognition, the sandwich's "filling" (Hurley, 2001). Traditional views of cognitive



(A)

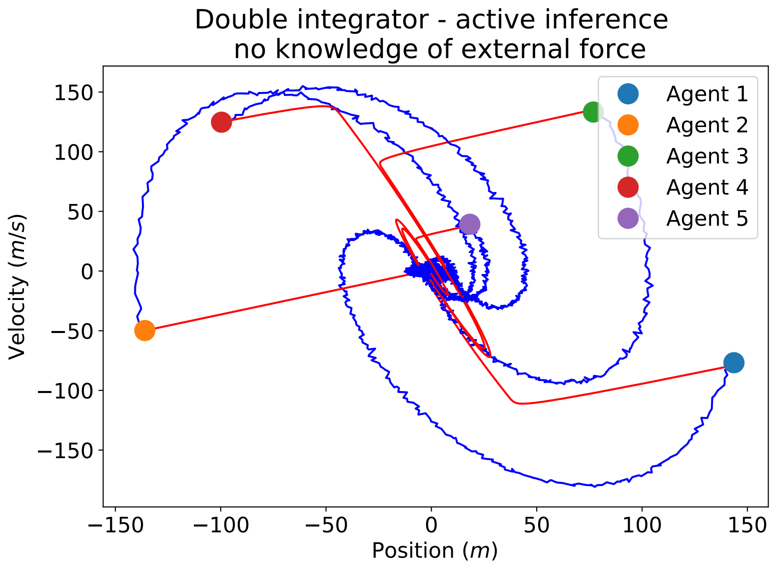


(B)

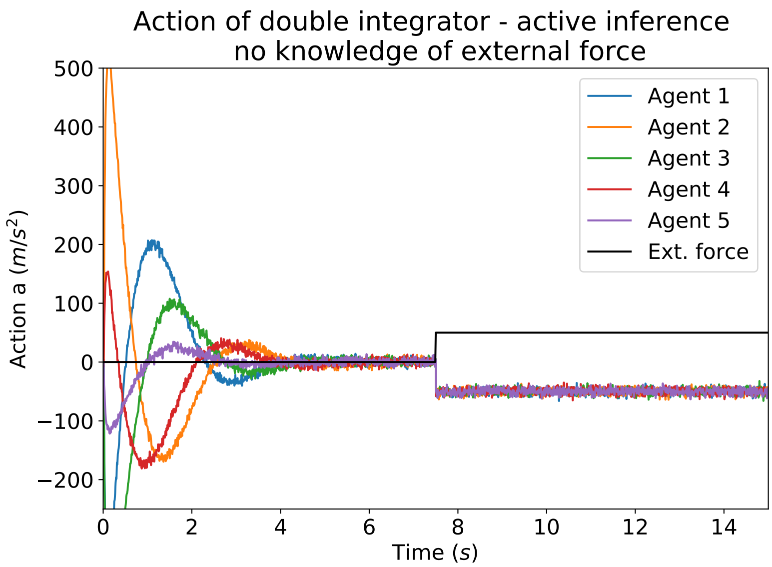
FIGURE 7.10: **The double integrator solved using active inference ($\alpha_1 = \exp(-1), \alpha_2 = \exp(0)$).** (a) Five active inference agents with different initial conditions and different parameters of the imaginary mass-spring system implemented as a prior, $\alpha_1 = \exp(-1), \alpha_2 = \exp(0)$. Showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents.

science openly embrace this architecture and the idea of modularity of perception and action while 4E theories largely reject them, claiming that fast-paced dynamic interactions between an agent and its environment imply that perception and action are deeply entangled and therefore not modular since such dynamics cannot be internally modelled (Varela, Thompson, and Rosch, 1991; Clark, 1998; Wilson, 2002; Di Paolo, Buhrmann, and Barandiaran, 2017).

In active inference, the more established modular role of action and perception



(A)



(B)

FIGURE 7.11: **The double integrator solved using active inference with no knowledge of external forces.** Same layout as Fig. 7.4. (a) Five active inference agents with different initial conditions showing the observed trajectories of the blocks in the phase-space (in blue) and the agents' estimates of the trajectories (in red). (b) Actions taken by the five agents (coloured lines) and external force applied to the agent (in black).

has already been put in discussion (Friston, 2011; Pickering and Clark, 2014; Seth, 2014b; Engel, Friston, and Kragic, 2016; Wiese, 2016; Pezzulo et al., 2017). We thus argued that in this framework not only action and perception are not modular in a Fodorian (Fodor, 1983; Coltheart, 1999; Prinz, 2006) or post-Fodorian (Barrett and Kurzban, 2006) sense, but that their intimate relation challenges our common understanding of these processes per se. To support this claim, we then connected the idea of modularity in cognitive science to concepts from control theory in order to define it using established mathematical frameworks. The concept of modularity itself is often considered vague (Coltheart, 1999; Barrett and Kurzban, 2006), Fodor himself defined modularity only *"to some interesting extent"*, (Fodor, 1983). However, we believe that information encapsulation (Fodor, 1983) should be considered a necessary requirement for this idea (as pointed out by Fodor more recently (Fodor, 2001), see also Prinz (2006)) , highlighting (1) restricted access to higher order information and vice versa, and (2) information content limited and specific to each module. We then suggest that the concept of modularity can be constructively formalised using the *separation principle* of control theory (Wonham, 1968; Åström and Murray, 2010; Georgiou and Lindquist, 2013). This proposal follows a general trend in the fields of cognitive science, neuroscience and psychology to use theories of estimation and control to account for perception and action respectively (Knill and Richards, 1996; Rao and Ballard, 1999; Kawato, 1999; Todorov, 2004; Todorov, 2009b; Franklin and Wolpert, 2011). All these proposals, however, suggest (often implicitly) that the two processes can be treated separately. On the other hand, 4E formulations normally assume that the presence of feedback mechanisms in tightly coupled agent-environment systems is indicative of non-modular systems (Beer, 1995; Clark, 1998; Hurley, 2001; Di Paolo and Iizuka, 2008; Di Paolo, Buhrmann, and Barandiaran, 2017; Wilson and Foglia, 2017).

With the architecture defined by the separation principle, i.e., the Linear Quadratic Gaussian (LQG) control framework, it is however possible to implement agent-environment systems where a separation of estimation (perception) and control (action) for an agent is not only possible, but optimal according to a specific list of criteria. While still focusing on action and perception, it is also important to highlight that following the definition of "learning" given in Chapter 3 and based on Friston (2008a), the conditions for this separation further introduce computationally tractable ways of learning the parameters of a (linear) dynamical system, a process also known as *system identification*. For instance, using methods based on the standard Ho-Kalman algorithm (Ho and Kalman, 1966), or its stochastic versions (Akaike, 1974; Van Overschee and De Moor, 1993), and given the *Markov parameters* of a system (i.e., its impulse response), one can identify an Hankel (block) matrix encoding conditions of observability and controllability, and subsequently recover state transition, input-state and observation mappings of a system, described by matrices A, B and C. These requirements are, nevertheless, very restrictive and possibly unrealistic for the study of real-world organisms. To obviate these limitations, the active inference formulation of sensorimotor loops (Friston et al., 2010a; Friston, 2010b; Buckley et al., 2017) proposes an extension of the traditional LQG-based architectures that does not implement the assumptions required by the separation principle. Following our claim regarding the modularity of the mind as an analogy of this principle, we thus claim that active inference aligns in this sense with themes closer to 4E cognition, and in particular formalises some of the implicit claims made by 4E researchers.

7.5.1 LQG vs Active inference, different mathematical formulations

The framework provided by LQG control linearly combines two processes of 1) estimation or inference of hidden properties of the environment and 2) control or regulations of variables of interest. The estimation of hidden variables is based on the presence of a Kalman (for discrete time systems) or Kalman-Bucy (for continuous time systems) filter, while the control of the desired variables on a Linear Quadratic Regulator (LQR) (Wonham, 1968; Anderson and Moore, 1990; Stengel, 1994). In particular, this combination is provably optimal according to a set of specifications:

1. the estimator must be implemented through a state-space model where only linear process dynamics and observation laws describe the environment and its latent variables

2. uncertainty or noise processes in both dynamics and observations are represented by white, zero-mean Gaussian variables

3. the properties of these random variables, and in particular their (co)variance matrices representing uncertainty are known

4. the performance of the regulator can be evaluated using a quadratic cost function

5. all the inputs to the agent are known, e.g., external disturbances and internal bodily signals such as motor actions.

As we argued in section 7.3.1, the first four criteria could be considered reasonable approximations of real-world phenomena in some cases, but the last one is more problematic. One of the main features of living systems is in fact the constant and inherent lack of knowledge of exogenous input signals, or forces, from the world (Sontag, 2003). Moreover, these inputs include also motor signals thought to be essentially discounted from the set of real observations of an agent in order for the agent to differentiate between exafferent and reafferent input (Holst and Mittelstaedt, 1950; Crapse and Sommer, 2008), i.e., forces coming from the world or initiated by the organism. In the state-space representation advocated by LQG, both these classes of inputs are included under variables $\boldsymbol{u}$, that we defined as a superset of motor actions $\boldsymbol{a}$ in equation (7.11) and external forces $\boldsymbol{I}$.

In our active inference implementation, the controller is built using a generative model that still represents the agent as a block but now includes beliefs about the dynamics of the system as if the block was attached to a spring and a damper (e.g., a piston). The fictitious spring stabilises the position of the agent around its desired position, while the imaginary friction generated by the damper allows it to slow down and eventually stop, i.e., at velocity zero. The target trajectory, i.e., position = velocity = 0, is specified using prior expectations encoded by the agent. The active inference solution of the problem takes thus a radically different stance on problems of regulation. In traditional formulations based on forward and inverse models (Kawato, 1999; Wolpert and Ghahramani, 2000), the control strongly relies on the presence of an internal model that inverts the causal chain of actions generating observations in order to estimate the best action or policy that can bring about a desired observed state. In active inference, on the other hand, action is defined using an extrinsic frame of reference, combining a Bayesian model inversion scheme on proprioceptive sensations with innate reflex arcs minimising (proprioceptive) prediction errors through actions on the world (see section 3.2.2).

Active inference agents require direct observations of proprioceptors that can be used to trigger reflexes once (proprioceptive) prediction errors are generated. In our example, this means that agents need access to observations on their acceleration since their motor output is a force. In standard LQG architectures, this information can be seen as intrinsic, since the Kalman-Bucy filter needs full knowledge of inputs $\boldsymbol{a}$ (and external ones $\boldsymbol{I}$ if present) to appropriately estimate hidden states $\boldsymbol{x}$ (Stengel, 1994). The active inference version may look more demanding because derivatives of noisy readings are usually to be avoided, but in standard implementations of Kalman-Bucy filters as models of perception, it remains unclear if a copy of motor outputs $\boldsymbol{a}$ can even be used at all, see discussions on efference copy in Feldman et al. (2015), Feldman (2016), and Friston (2011). In support of the active inference proposal, it has been shown that even unicellular organisms can use simple mechanisms estimating derivatives of their observations if they have no dedicated receptors, e.g., temporal sensing of chemical concentration gradients in E. Coli (Andrews, Yi, and Iglesias, 2006). Furthermore, by eschewing explicit internal representations of external inputs and including observations on acceleration, active inference improves the controller used for problems of regulation, now necessarily including an integral-like component (see Chapter 6). The presence of this extra term can easily be explained by the known property of integral controllers to approximate via a linear mechanism the presence of unknown inputs (Sontag, 2003). In this context, integral control can be seen as the "optimal" linear controller in the presence of uncertainty, i.e., when input signals (motor commands or external forces) are not known.

A crucial difference between LQG and active inference lies in the definition of the Kalman gain, K, and feedback gain, L, matrices. These two matrices essentially prescribe the optimal update speed of estimates of hidden states and controls. Both of them require solving Riccati equations involving knowledge of the covariance matrices of dynamics and observation noise in the first case, and weights representing costs for estimation and control for the latter. However, K and L are in general not well defined in active inference since their formulation requires the use of Markovian stochastic processes (Stengel, 1994). These processes, essentially implemented using Wiener noise, are often considered a good approximation of coloured noise when the time scale of the estimator is much slower than the one of the true dynamics of the environment. However, this hardly holds in practical implementations in engineering with a high sampling rate (Stratonovich, 1967). Similarly, biological systems have a characteristic time scale closer to real dynamics, implying that estimation mechanisms must somehow be dealing with non-Markovian dynamics (Friston, 2008a). The apparent complication derived from the definition of the so called *generalised coordinates of motion*, implementing generalised state space models where non-Markovian, analytical stochastic processes are represented becomes in this sense a small token to pay. Furthermore, it can be seen as necessary in cases where the underlying stochastic dynamics are themselves non-Markovian processes, as pointed our previously for applications in engineering (Stratonovich, 1967; Jazwinski, 1970), physics (Fox, 1987; Van Kampen, 1992; Łuczka, 2005) and neuroscience (Friston, 2008a; Valdes-Sosa et al., 2011; Li et al., 2011). In active inference, the Kalman gain, K, and feedback gain, L, matrices are replaced by learning rates such as in this work or in the implementation presented in Chapter 5, or by clever implementations that allow for adaptive update schemes with varying integration steps as in (Friston, Trujillo-Barreto, and Daunizeau, 2008), see also Chapter 4.

Another point of contention can be found in the cost function(al) minimised in the two cases: a value or cost-to-go function for LQG, equation (7.5), and a variational free energy functional for active inference, equation (3.8). As we can see, the free energy functional presents more terms than the value function of LQG. The presence of multiple prediction errors in the variational free energy formulation can be explained by the fact that errors map directly to likelihood and prior distributions as obtained from the joint density of observed and hidden states used to define free energy, see equation (3.33). In LQG, on the other hand, there's an often implicit assumption on the invertibility of matrix C, mapping estimated states $\hat{x}$ to observations $\boldsymbol{y}$, see eqn. 4.1-1 to 4.1-8 in (Anderson and Moore, 1990) for a clarification. Thanks to this assumption, the desired/target trajectory can be expressed directly into the frame of reference of hidden, rather than measured states. While this may be a fair assumption for many systems in LQG, solutions for non-invertible matrices, i.e., as in the more general nonlinear case, are still not available. In active inference, a pseudo-inversion is provided through Bayes theorem and approximated with a variational formulation that can encompass also nonlinear systems (Friston, Trujillo-Barreto, and Daunizeau, 2008; Buckley et al., 2017).

7.5.2 LQG vs Active inference, repercussions for the cognitive sciences

LQG-based architectures are modular in nature, with perception and action seen as separate problems solved nearly independently (Kawato, 1999; Wolpert and Ghahramani, 2000; Todorov, 2004). According to this view, initially a system should find accurate estimates of the hidden properties of its observations, and only once such estimates are available should an agent attempt to regulate variables that are of interest to achieve its goals, e.g., temperature, oxygen level, etc.. On the other hand, we can define a framework based on mathematical formulations of control problems where the separation principle is not included or required. According to one such proposal, that we identified in active inference (Friston et al., 2010a; Friston et al., 2010b; Buckley et al., 2017), perception and action are combined in an inseparable sensorimotor loop described by the minimisation of variational free energy for an agent. In this set up, action and perception are seen as instances of a fundamentally unique functional process (Clark, 1998), using different labels for our (i.e., the observers') convenience. In particular, the idea of precise inferences of world variables is called into question (Clark, 2015a), to the point that inaccurate perception is not only possible but becomes a pre-requisite to act on the world (Brown et al., 2013; Wiese, 2016). In architectures based on the separation principle, the estimated state of a system is thought of as a relevant account of real observations, e.g., their means and covariances. Conversely, in active inference it becomes clear that estimates of latent variables of the world are deeply connected to the current goal of an agent, e.g., to regulate its observations, cf. (Powers, 1973). To do so, its targets are encoded as prior expectations and used to bias inferential processes toward its desires so that prediction errors are created as the mismatch of observations and the estimates of hidden variables. These errors are then minimised by acting on the world (Friston et al., 2010a), taking advantage of proprioceptive prediction errors that enact reflex arcs to make observation better accord with existing predictions (Clark, 2015b; Wiese, 2016). More in general, the active inference formulation allows also for accurate estimates of the latent variables generating observations, see for instance (Friston, Trujillo-Barreto, and Daunizeau, 2008), but this modality fundamentally excludes the possibility of acting: if no prediction errors are generated for action to minimise, an agent becomes a simple mirror of its world with no strong desire or even necessity to act (Friston, Thornton, and Clark, 2012; Brown et al., 2013) (see also Chapter 4). In other words, depending on different precision weights an agent can accurately estimate its observations without acting or potentially discard its sensations to only pursue its desires, generating all possible cases in between as a balanced mix of weighted prediction errors (Allen and Friston, 2018), see also Chapter 4.

In recent years, the more traditional understanding of perceptual and motor as nearly independent processes as been put into discussion by different authors especially in neuroscience (Ahissar and Assa, 2016; Busse et al., 2017; Buckley and Toyoizumi, 2018). It appears clear that many experimental set ups are limited in a way that decisively affects conclusions that claim or just presuppose the separation of perceptual and behavioural components (Krakauer et al., 2017), requiring new and ethologically meaningful paradigms for an appropriate study of different aspects of living systems (Najafi and Churchland, 2018). In this context, we make an attempt to propose some new ideas that could be used to drive future experiments. This attempt is centred around a critical appraisal of LQG as a model architecture for cognitive systems, focusing in particular on the assumptions made by the use of Kalman-Bucy filters, central to these proposals (Wolpert, Ghahramani, and Jordan, 1995; Todorov and Jordan, 2002; Wolpert, Diedrichsen, and Flanagan, 2011;

Franklin and Wolpert, 2011). One of the key requirements for Kalman-Bucy filters to generate an accurate estimate of the hidden state of a system is to have access to both *all* the outputs (the observations) and the *all* inputs (forces that affect the state of a system). Such inputs include both motor commands, which in classical forward/inverse models are identified using the idea of efference copy (Holst and Mittelstaedt, 1950) (see for instance Kawato (1999), Wolpert and Ghahramani (2000), and Todorov (2004)), but also external forces/signals from the environment that cannot be accounted by an organism, i.e., a sudden change in weather conditions or unexpected interactions with other agents.

The presence of the former has amply been put into discussion for decades, mainly through the proposal of the equilibrium point hypothesis or its most recent version, threshold or referent control (see some recent discussions in Feldman (2009), Feldman et al. (2015), and Feldman (2016)). Supported by neurophysiological arguments, the main proponents of this approach strongly advocate for a theory of motor control that eschews the idea of efference copy, deemed to be inconsistent with experimental results (Feldman, 2016). In its place, referent control proposes moving frames of reference that implement different equilibrium points, or trajectories in more general terms, essentially shifting muscle activation thresholds based on different desired sensory consequences. While inverse models (Kawato, 1999) require motor commands to be defined in an intrinsic frame of reference created through the inversion of a forward model encoding cause-effect relationships accurately representing the physical laws underlying observations, referent control essentially works at the periphery of an organism via the use of reflex arcs triggering different muscle activations. Active inference largely follows arguments from threshold control, explaining motor behaviour using a refined version of this hypothesis including a mathematical formulation based on predictive coding (Adams, Shipp, and Friston, 2013). Agents (or their body parts, limbs, joints) are thus represented as mass-spring systems whose equilibrium encodes a desired trajectory in the phase-space, see our example or Friston et al. (2012). The debate over the role of efference copy remains largely unsettled (Crapse and Sommer, 2008; Wolpert, Diedrichsen, and Flanagan, 2011; Feldman, 2016; Schwartz, 2016; Straka, Simmers, and Chagnaud, 2018) and new investigations are necessary to definitely explain its importance. In this context, new experiments to disambiguate the role of efference copy are necessary, requiring also a deep clarification of the terminology, often still confusing (cf. efference copy vs. corollary discharge (Crapse and Sommer, 2008; Schwartz, 2016; Straka, Simmers, and Chagnaud, 2018)).

The presence of external unaccounted forces is often overlooked in many experimental set-ups with fixed or predictable conditions (e.g., the classic and still dominating two-alternative forced choice paradigm). In more realistic and ethological scenarios, however, one should expect that external and unpredictable stimuli constantly affect the behaviour of an agent (Krakauer et al., 2017; Najafi and Churchland, 2018; Buckley and Toyoizumi, 2018). In this case, introducing (non-zero-mean) noise or varying experimental conditions may help in testing the robustness of LQG-based architectures. In practice if some inputs are not known, one should expect LQG to perform rather poorly until these inputs can be estimated and adaptation (e.g., learning) to new conditions can take place. However, one should then explain how such forces can be estimated in LQG since Kalman-Bucy filters cannot estimate inputs (Kalman and Bucy, 1961; Chen, 2003), cf. DEM (Friston, Trujillo-Barreto, and Daunizeau, 2008). More in general, if a system is instead well adapted to deal with unpredictable stimuli, simple mechanisms such as integral control could be in place, as shown formally in Sontag (2003) and in experiments on chemotactic adaptation in E. Coli (Yi et al., 2000). More recently, some promising results were presented in Ritz et al. (2018), supporting the idea that integral feedback control is a good model for adaptation in environments with varying conditions. Integral control constitutes the best linear approximation to problems of control with unknown forces affecting the observations of an agent (Åström, 1995), providing a robust solution with fast responses to problems that otherwise would require slower learning mechanisms (Yi et al., 2000), which may be ineffective in fast-paced environments (Ashby, 1958).

7.6 Conclusions

In this work we proposed to ground a long-standing debate over the modularity of action and perception (Fodor, 1983; Coltheart, 1999; Barrett and Kurzban, 2006) that characterises some aspects of the feud between traditional and 4E accounts of cognition (Van Gelder, 1995; Clark, 1998; Hurley, 2001; Wilson and Foglia, 2017) using the separation principle from control theory (Wonham, 1968; Anderson and Moore, 1990; Stengel, 1994). The notion of modularity implies a formal characterisation of perception and action as distinct and encapsulated processes, with limited interactions mediated by higher order cognition in a purely sequential fashion. Other definitions of modularity are often used to describe different physiological or functional units within perceptual (Wolpert, 1997) or motor apparatuses (Flanders, 2011; Zelik et al., 2014), but here we emphasised modularity as the functional segregation affecting perception and action more in general. In control theory, the separation principle advocates the optimal decomposition of estimation and control of a system following a set of assumptions entailing mainly the linearity of the system to control and full knowledge of its parameters and inputs.

In recent developments of cognitive science, estimation and control theory have been proposed to provide mathematical descriptions of several hypotheses on the nature of perception and action respectively (Knill and Richards, 1996; Knill and Pouget, 2004; Doya, 2007; Körding and Wolpert, 2006; Wolpert, Ghahramani, and Jordan, 1995; Kawato, 1999; Friston, 2010b). Our proposal thus entails extending this formal description of cognitive processes by using more advanced notions from control theory, in an attempt to further account for other ideas highly debated in the cognitive science community. It is interesting to note that the notion of separation emerged also in econometrics, in parallel to control theory, as "uncertainty equivalence" (Simon, 1956; Theil, 1957). The name is derived from the idea that the regulation of a stochastic system can be achieved using a deterministic controller after an estimation of the uncertain, hidden properties of the system (Stengel, 1994). One of its main proponents, Herbert Simon, is also one of the fathers of (symbolic) AI, known for its positions on near-decomposability (Simon and Ando, 1961; Simon, 1991), an idea describing dynamical systems (physical, biological, social) as hierarchical organisations of nearly independent modules. While no direct connection to Fodorian views has been made since Fodor focused more on aspects of the mind, different authors have identified near decomposability as a concept related to, if not indirectly inspiring the mainstream theory of the modularity of the mind (Velichkovsky, 2005; Bechtel, 2009).

In this light, we claimed that cognitivist/computational ideas based on the modular architecture can be seen in analogy with the separation principle and in doing so, we explicitly gave a list of requirements for the separation of action and perception, the ones used for the definition of the separation principle. This definition is however extremely strict. The principle can be applied only to a small subset of systems (linear, with Gaussian noise, quadratic cost functions, known covariances and known inputs). It is thus hard to imagine, on this view, how studies of brains and minds could make such assumptions, suggesting then that non-modular 4E views provide a more suitable framework for investigating cognitive and natural systems. We then argued in favour of a recent proposal based on theories of estimation (or inference) and control and with no explicit assumption regarding their separability: active inference. In this framework, one of the five necessary requirements for separability, the idea of having access to all inputs and in particular to a copy of motor signals and to (unexpected) stimuli from the environment, is dismissed (Friston et al., 2010a; Friston, 2011; Adams, Shipp, and Friston, 2013). By rejecting such mechanism, active inference effectively challenges classical architectures based on the separation principle and in doing so, we claimed, explicitly agrees with 4E views of cognition whereby perception and action are seen as non-modular processes. Our work provides thus support for hypotheses highlighting how active inference is more in agreement with 4E theories than with traditional accounts of cognition (Clark, 2015b; Bruineberg, Kiverstein, and Rietveld, 2018; Pezzulo et al., 2017).

As a final remark, to quote Kalman on an early intuition regarding the problem of simultaneous estimation and control (perception and action in our interpretation):

> One may separate the problem of physical realization into two stages: computation of the "best approximation" $\hat{x}(t_1)$ of the state from knowledge of $y(t)$ for $t \leq t_1$ and computation of $u(t_1)$ given $\hat{x}(t_1)$.

(Kalman, 1960c)

This may true for engineering but perhaps not for studies of cognition and natural systems.

Chapter 8

Conclusions

This thesis constitutes an attempt to ground ideas developed under the free energy principle (FEP) into embodied, enactive, embedded and extended (4E) cognitive science using the mathematical frameworks of probability theory and dynamical systems/control theory. In particular I focused on active inference, one of the most promising process theories implementing the minimisation of variational free energy proposed by the FEP. This minimisation is claimed to provide a general principle of self-organisation in theoretical biology (Friston, 2012; Friston, 2013), and accouts of perception, learning and action (among others) in neuroscience (Friston, Kilner, and Harrison, 2006; Friston, 2010b). While raising a lot of interest in different fields of the natural and cognitive sciences, several aspects of the FEP have remained obscure, with a proliferation of reviews by different authors (Bogacz, 2017; Buckley et al., 2017; Biehl et al., 2018) attempting to shed light on some of the ambitious claims made by the FEP. This thesis differs from work by Bogacz (2017) due to my focus on sensorimotor loops rather than purely perceptual accounts of cognitive systems. Chapter 3 builds on Buckley et al. (2017), which develops a formal mathematical derivation of the FEP in generalised coordinates of motion for non-Markovian univariate continuous time stochastic processes in a hierarchical set up. The following chapters however cover specific connections of the FEP to 4E theories of cognitive science using a series of agent-based simulations to support my claims while showing some of the more explicit connections to the field of control theory. In Biehl et al. (2018) we find a general treatment of different theories of intrinsic motivation: *expected* variational free energy minimisation, predictive information, empowerment and knowledge seeking, with formal connections of the FEP to the universal reinforcement learning framework. My focus is however different, and intrinsic motivation proposals are never fully discussed. Moreover, this thesis adopts a different formulation of the FEP (minimisation of variational free energy vs minimisation of expected variational free energy), a different mathematical description of state-space models (continuous time vs discrete time/(PO)MDPs), and has a different aim (connections of the FEP to cognitive science vs a unified formalisation of different mathematical frameworks and a discussion of intrinsic motivation theories).

Using different agent-based models I provided:

1. a series of experiments clarifying the implementation of the FEP for minimal models of sensorimotor loops, with
2. concrete connections between 4E's theories of cognition and the notion of variational free energy minimisation formalised through Bayesian inference and control theory.

Specifically, in Chapter 4 I implemented an agent model that captures the main features of the active inference proposal, focusing on the importance of action for Bayesian models of cognition that otherwise often revolve around only perceptual processes. In these simulations I also introduced two of the main constructs of this process theory: precision hyperparameters, a set of weights, or gains, regulating the minimisation of different prediction errors and priors, a set of probability distributions (or beliefs) encoding subjective preferences or credences of an agent. In doing so I highlighted how, in order to trigger purposeful behaviour in an active inference agent, the common and intuitive idea of correctly inferring properties of the world is in fact misleading. Only by having high precisions on priors that are misaligned with world states can an agent generate (proprioceptive) prediction errors for action to resolve, generating what appears as purposeful behaviour (Wiese, 2016).

In Chapter 5 I built on this first example and proposed that generative models representing active inference agents need not represent any objective property of their milieu. Priors that create a strong coupling between and agent and its environment can, in principle, be sufficient to enact complex behaviours as long as they encode appropriate sensorimotor contingencies for the agent.

Chapter 6 proposed a general-purpose methodology following the idea of simplified generative models introduced in the previous chapter, implementing a PID-like controller in active inference. PID control is one of the most popular strategies used by modern controllers in industry since it constitutes a simple but effective mechanism of robust regulation in presence of external disturbances and noise. In active inference, this method is equivalent to an "agent" implementing a generative model in terms of first order linear (stochastic) differential equations. Strong priors on a desired set trajectory to follow generate then prediction errors that the agent tries to minimise by producing actions that can make the world behave according to its target, achieving regulation by imposing its desired dynamics on the world. The relevance of this formulation can also provide a direct connection between the FEP and models of PID control used in biology and neuroscience to model different behaviours, e.g., chemotaxis in E. Coli (Yi et al., 2000) or adaptation to unexpected stimuli in psychophysics (Ritz et al., 2018).

In Chapter 7 I discussed in detail one of the fundamental ideas behind the implementation of active inference: the combination of perception and action into a unified process, challenging the more traditional idea of modularity between the two (Fodor, 1983; Hurley, 2001). This idea has been inspirational for computational theories of the mind, with a central tenet regarding perception and action as separate, informationally encapsulated functional modules. Perception and action are nowadays often described using estimation and control theory, upon which the FEP and other frameworks are based. To try to formalise modularity I used the separation principle of control theory, asserting that the estimation of a system and its regulation can be solved as independent problems and sequentially implemented for the control of partially observable systems. Using the separation principle I thus proposed a mathematical connection between non-separable implementations of sensorimotor loops in active inference and 4E theories advocating for deeply entangled perceptual and motor processes (Clark, 1998; Wilson, 2002).

8.1 The FEP - promises and deliverables

The FEP has attracted a lot of attention for its ambitious claims regarding the unification of different mathematical frameworks (information, dynamical systems and control theory) for the study of living organisms in the natural and cognitive sciences (cognitive (neuro)science and biology) based on arguments from the physical sciences (thermodynamics and statistical mechanics).

The combination of ideas from information theory and dynamics systems/(feedback) control was proposed in the last century with influential insights and contributions from Ashby (1958), Shannon (1959) and Kalman (1960c), among others. New developments have since then improved our understanding of the fundamental relationships between these frameworks, showing the fundamental duality of the problems of control and estimation initially formalised by Kalman (1960c), see for instance Attias (2003), Mitter and Newton (2003), Todorov (2008), and Kappen, Gómez, and Opper (2012). See especially Tishby and Polani (2011) and Beer and Williams (2015) for discussions related to action-perception loops. In my work, I provided an extension of this combination in Chapter 6, where the control theoretical formulation of PID control was derived from the minimisation of variational free energy for simple linear generative models. This implementation has also introduced an analytical method to find the optimal parameters of PID controllers as a slower process of free energy minimisation where, more traditionally, only heuristics and trial-and-errors procedures exist.

Estimation and control have been used to account for perception and action in cognitive and motor (neuro)science, becoming the dominant analogies and computational/algorithmic theories for these cognitive functions, e.g., Knill and Richards (1996), Rao and Ballard (1999), Lee and Mumford (2003), Kawato (1999), Wolpert and Ghahramani (2000), and Todorov (2004). These theories however still heavily rely on linear models where perception and action have similar implementations (Kalman, 1960c; Todorov, 2008). The FEP builds on these proposals, already providing mathematical hypotheses on the nature of cognitive functions (Todorov, 2009b), by reinterpreting the role of action and behaviour using evidence from neurophysiology supporting an even more unified vision of sensorimotor loops (Friston et al., 2010a; Friston, 2011; Adams, Shipp, and Friston, 2013). In this light, perception and action are not only seen as co-dependent processes in agent-environment coupled systems, they are thought to be more deeply intertwined if not almost undistinguishable in some important aspects (Pickering and Clark, 2014; Wiese, 2016; Pezzulo et al., 2017). This notion differs from the idea of active sensing (Yang, Wolpert, and Lengyel, 2016), relying on an intrinsic drive for an agent's actions to gather more information in order to improve its perception of the world, and promotes regulation of an agent's variables (Ashby, 1960; Powers, 1973; Seth, 2014b; Engel, Friston, and Kragic, 2016) as a better interpretation of its behaviour:

> Its first characteristic is that its ultimate aim is not understanding but the purely practical one of control. If a system is too complex to be understood, it may nevertheless still be controllable. For to achieve this, all that the controller wants to find is some action that gives an acceptable result; he is concerned only with what happens, not with why it happens.

Ashby (1958)
(on operational research, although in my opinion this quote can also be seen as representing his view on biological organisms as control systems)

These ideas strongly resonate with some of the core themes from 4E cognitive science: agent-environment coupled systems and the importance of feedback, cognitive offloading to body and environment and an emphasis on control and behaviour over accurate descriptions and representations of the environment expressed in different chapters of this thesis.

In this work, I largely overlooked the potential relationships between the FEP and thermodynamics/statistical mechanics. These relationships are thought to be based on the conceptual arguments regarding the state of low (thermodynamic) entropy of living organisms (Friston, 2010b) as famously suggested by Schrödinger (1944). Furthermore, more formal connections have been proposed with the Maximum Entropy principle by Jaynes (Jaynes, 1957a; Jaynes, 1957b), see for instance (Friston, 2012). In both cases, the FEP relies on the relation between information and its thermodynamic costs (Parrondo, Horowitz, and Sagawa, 2015; Kolchinsky and Wolpert, 2018), and in particular between Shannon (information) and thermodynamic entropy (Sengupta, Stemmler, and Friston, 2013; De Ridder, Vanneste, and Freeman, 2014). The ambiguous use of "free energy" however, mostly derived by work in machine learning (Hinton and Zemel, 1994), often suggests a physical relationship between variational free energy and Helmholtz free energy which has not been formally shown[1]. In support of a possible connection between the two, Still et al. (2012) showed that, under a set of assumptions, increasing the predictive power

[1]On the other hand, Ramstead, Badcock, and Friston (2017) claim that an indirect relationship exists, based on the idea that minimising variational free energy, on average, minimises thermodynamic entropy.

of a system from an information theoretical perspective is equivalent to minimising Helmholtz (thermodynamic) free energy. The assumptions are however quite restrictive, mainly a system having no feedback from the environment, and the predictive power is measured using a different definition from the one used by the FEP (i.e., model evidence, minimising surprisal is equivalent to maximising model evidence, see equation (3.4)). Further connections may be found in the future, considering work on the thermodynamics of feedback control, see Kolchinsky and Wolpert (2018) and references therein, clarifying for instance the connections between different information measures of predictive power and their implications (Sengupta, Stemmler, and Friston, 2013). More generally, this aspect of the FEP remains largely unexplored, and will need a thorough formalisation based on stronger arguments for non-equilibrium thermodynamics describing biological organisms[2].

8.2 Ideas for the future, Bayesianism and ways forward

The use of Bayesian methods for the analysis of data is nowadays ubiquitous (Rahnev and Denison, 2018). More interestingly, as I've argued in this thesis, these methods are nowadays used as analogies of different cognitive functions. Bayesian reasoning offers a fresh perspective on different aspects of decision making (Robert, 2007) and is even claimed to be an extension of mathematical logic (Jaynes, 2003). In this section I'll discuss two potential problems with the adoption of Bayesian reasoning that one should consider to gain a proper understanding of my work, suggesting ways to interpret my contributions and possible future directions I'd like to explore.

One of the major risks is to over-generalise the use of Bayesian methods to describe different natural and cognitive phenomena, e.g., Bayesian inference in plants? (Calvo and Friston, 2017). This often implies the presence of a strong form of *Bayesianism* (Pearl, 2001), a tendency to claim "Bayes-optimal" computation here, there and everywhere (although nowadays an increasing number studies attempt to compare Bayesian and non-Bayesian, i.e., not reliant on sensory/process uncertainties, models (Adler and Ma, 2018; Acerbi et al., 2018)). This tendency is usually due to the overly "generous" definition of Bayesian optimal processes and mainly to the presence of *subjective* priors (Jones and Love, 2011; Rahnev and Denison, 2018). It is beyond the scope of my work to dive into a deep philosophical debate on the value of this and possibly other constructs in the Bayesian formalism (see Jaynes (2003) for some biased, but thorough discussion), but I'll briefly discuss a few ideas explaining my pragmatism in this case. This thesis may appear as a canonical example of the Bayesianism mentioned above and the reader may ask for example: 1) what is the reason to implement a simple Braitenberg vehicle using the complicated FEP framework (Chapter 5)? Or: 2) what can you gain from rewriting a PID control in

[2] At the same time I am writing this paragraph I attended a lab meeting in Friston's group, where the main discussion focused on a tentative and more coherent, unified description of biological organisms from the perspectives of random dynamical systems, classical mechanics, information theory and geometry, and especially non-equilibrium thermodynamics.

terms of approximate Bayesian inference (Chapter 6)? In the first case, the rationale was to extend existing pedagogical implementations of the FEP such as McGregor, Baltieri, and Buckley (2015), Bogacz (2017), and Buckley et al. (2017) to topics of 4E cognitive science. This example introduced a (hopefully) clearer perspective on embodiment and the importance of situatedness and dynamic coupling over complex models of the world within a Bayesian contest with parsimonious generative models (Seth, 2014b; Clark, 2015a). To answer the second question, it is important to highlight the potential of recasting problems of control in terms of Bayesian inference (Doya, 2009; Friston, 2011), with prominent examples including for instance Todorov (2009a) and Kappen, Gómez, and Opper (2012). The formulation I provide in Chapter 6 includes, for instance, an analytical method for finding the parameters of a PID controller based on the presence of simple priors. The tuning of these parameters usually requires ad-hoc trial-and-error procedures that cannot generalise, as different books and chapters dedicated to such heuristics show (Åström, 1995; Johnson and Moradi, 2005; Åström and Hägglund, 2006).

The second issue emphasises the "evidence-consistent-with"-problem[3]. As with many other statistical accounts, most of the work presented under the predictive coding/processing, Bayesian brain and FEP frameworks is correlational in nature (i.e., the measures used can only describe correlations/associations), and supported at the moment only by "evidence consistent with" these hypotheses. This has strong repercussions on the idea that agents might be engaging in processes of Bayesian inference. A major source of concern may thus emerge for scientists interested in, for instance, counterfactual reasoning (Seth, 2014a). Even without considering the issues with the FEP and the unimodal Laplace (Gaussian) assumption (Seth, 2014b), not defining multiple outcomes and tackled with the *expected* free energy minimisation framework (Friston, Samothrakis, and Montague, 2012a; Friston et al., 2015), a potential risk is to think that the problem of counterfactuals is solved just with the presence of spatially and temporally deep generative models (Friston et al., 2017; Kirchhoff et al., 2018). These models lack the ability to retrospect and potentially other defining features of counterfactual thinking (Pearl and Mackenzie, 2018). For a better definition of counterfactuals we may instead wish to consider work on causal models extending the traditional language of (correlational) probability theory, for instance through the implementation of operations based on *do-calculus* (Pearl, 2009). This framework is claimed to provide a more appropriate venue for investigating counterfactuals and interventions on distributions (Pearl and Mackenzie, 2018). Some of the most recent formulations of the FEP (Friston, Parr, and Vries, 2017) suggest that the framework adopted by active inference theories is, in principle, compatible with the Bayesian network formulation of Bayesian networks in Pearl's causal modelling (Pearl, 2009), but further explorations will be necessary in the (near) future.

[3]This line emerged from several discussions with Warrick Roseboom.

Bibliography

Acerbi, L. et al. (2018). "Bayesian comparison of explicit and implicit causal inference strategies in multisensory heading perception". In: *PLoS computational biology* 14.7, e1006110.

Adams, R. A., Shipp, S., and Friston, K. J. (2013). "Predictions not commands: active inference in the motor system". In: *Brain Structure and Function* 218.3, pp. 611–643.

Adams, R. A. et al. (2013). "The Computational Anatomy of Psychosis". In: *Frontiers in Psychiatry* 4, p. 47.

Adler, W. T. and Ma, W. J. (2018). "Comparing Bayesian and non-Bayesian accounts of human confidence reports". In: *PLoS computational biology* 14.11, e1006572.

Ahissar, E. and Assa, E. (2016). "Perception as a closed-loop convergence process". In: *Elife* 5, e12830.

Ahmadi, A. and Tani, J. (2018). "Learning to Embed Probabilistic Structures Between Deterministic Chaos and Random Process in a Variational Bayes Predictive-Coding RNN". In: *arXiv preprint arXiv:1811.01339*.

Aitchison, L. and Lengyel, M. (2017). "With or without you: predictive coding and Bayesian inference in the brain". In: *Current opinion in neurobiology* 46, pp. 219–227.

Akaike, H. (1974). "Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes". In: *Annals of the Institute of Statistical Mathematics* 26.1, pp. 363–387.

Allen, M. and Friston, K. J. (2018). "From cognitivism to autopoiesis: towards a computational framework for the embodied mind". In: *Synthese* 195.6, pp. 2459–2482.

Amari, S. (2016). *Information geometry and its applications*. Springer.

Anderson, B. and Moore, J. B. (1990). *Optimal control: linear quadratic methods*. Prentice-Hall, Inc.

Andrews, B. W., Yi, T.-M., and Iglesias, P. A. (2006). "Optimal noise filtering in the chemotactic response of Escherichia coli". In: *PLoS Comput Biol* 2.11, e154.

Ang, J. et al. (2010). "Considerations for using integral feedback control to construct a perfectly adapting synthetic gene network". In: *Journal of theoretical biology* 266.4, pp. 723–738.

Ang, K. H., Chong, G., and Li, Y. (2005). "PID control system analysis, design, and technology". In: *IEEE transactions on control systems technology* 13.4, pp. 559–576.

Araki, M. and Taguchi, H. (2003). "Two-degree-of-freedom PID controllers". In: *International Journal of Control, Automation, and Systems* 1.4, pp. 401–411.

Arturo Urquizo (2011). *PID controller – Wikipedia, the free encyclopedia*. [Online; accessed March 30, 2018]. URL: https://commons.wikimedia.org/wiki/File:PID_en.svg.

Ashby, W. R. (1957). *An introduction to cybernetics*. Chapman & Hall Ltd.

— (1958). "Requisite variety and its implications for the control of complex systems". In: *Cybernetica* 1, pp. 83–99.

— (1960). *Design for a brain*. Wiley New York.

Åström, K. J. (1970). *Introduction to stochastic control theory*. Academic Press.

— (1995). *PID controllers: theory, design and tuning*. Research Triangle Park.

Åström, K. J. and Hägglund, T. (2001). "The future of PID control". In: *Control engineering practice* 9.11, pp. 1163–1175.

— (2004). "Revisiting the Ziegler–Nichols step response method for PID control". In: *Journal of process control* 14.6, pp. 635–650.

— (2006). *Advanced PID control*. ISA-The Instrumentation, Systems, and Automation Society Research Triangle Park, NC.

Åström, K. J. and Murray, R. M. (2010). *Feedback systems: an introduction for scientists and engineers*. Princeton university press.

Åström, K. J., Panagopoulos, H., and Hägglund, T. (1998). "Design of PI controllers based on non-convex optimization". In: *Automatica* 34.5, pp. 585–601.

Attias, H. (2003). "Planning by probabilistic inference." In: *AISTATS*. Citeseer.

Ay, N. et al. (2008). "Predictive information and explorative behavior of autonomous robots". In: *The European Physical Journal B* 63.3, pp. 329–339.

Ay, N. et al. (2012). "Information-driven self-organization: the dynamical system approach to autonomous robot behavior". In: *Theory in Biosciences* 131.3, pp. 161–179.

Azevedo-Filho, A. and Shachter, R. D. (1994). "Laplace's method approximations for probabilistic inference in belief networks with continuous variables". In: *Uncertainty Proceedings 1994*. Elsevier, pp. 28–36.

Baldeweg, T. (2006). "Repetition effects to sounds: evidence for predictive coding in the auditory system". In: *Trends in cognitive sciences* 10.3, pp. 93–93.

Baltieri, M. (2014). *A free energy principle for path tracking in a 1D world (MSc thesis)*.

Baltieri, M. and Buckley, C. L. (2017). "An active inference implementation of phototaxis". In: *Proc. Eur. Conf. on Artificial Life*, pp. 36–43.

— (2018a). "A Probabilistic Interpretation of PID Controllers Using Active Inference". In: *From Animals to Animats 15*. Ed. by P. Manoonpong et al. Springer International Publishing, pp. 15–26.

— (2018b). "The modularity of action and perception revisited using control theory and active inference". In: *The 2018 Conference on Artificial Life: A Hybrid of the European Conference on Artificial Life (ECAL) and the International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*. Ed. by T. Ikegami et al., pp. 121–128.

Baltieri, M. and Buckley, C. L. (2019a). "Modularity, the separation principle and active inference (in prep.)" In: –.

— (2019b). *Nonmodular architectures of cognitive systems based on active inference (submitted)*.

— (2019c). "PID control as a process of active inference with linear generative models (under review)". In: –.

Bar-Shalom, Y. and Tse, E. (1974). "Dual effect, certainty equivalence, and separation in stochastic control". In: *IEEE Transactions on Automatic Control* 19.5, pp. 494–500.

Barandiaran, X. E., Di Paolo, E. A., and Rohde, M. (2009). "Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action". In: *Adaptive Behavior* 17.5, pp. 367–386.

Barrett, H. C. and Kurzban, R. (2006). "Modularity in cognition: framing the debate". In: *Psychological review* 113.3, p. 628.

Barto, A., Mirolli, M., and Baldassarre, G. (2013). "Novelty or surprise?" In: *Frontiers in psychology* 4, p. 907.

Bastos, A. M. et al. (2012). "Canonical microcircuits for predictive coding". In: *Neuron* 76.4, pp. 695–711.

Baum, L. E. and Eagon, J. A. (1967). "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology". In: *Bulletin of the American Mathematical Society* 73.3, pp. 360–363.

Beal, M. J. (2003). *Variational algorithms for approximate Bayesian inference*. University of London London.

Bechtel, W. (2009). "Explanation: Mechanism, modularity, and situated cognition". In: *The Cambridge handbook of situated cognition*. Ed. by P. Robbins and M. Aydede. Cambridge University Press Cambridge, MA, pp. 155–170.

Beer, R. D. (1995). "A dynamical systems perspective on agent-environment interaction". In: *Artificial Intelligence* 72.1-2, pp. 173–215.

— (1997). "The dynamics of adaptive behavior: A research program". In: *Robotics and Autonomous Systems* 20.2-4, pp. 257–289.

— (2003). "The dynamics of active categorical perception in an evolved model agent". In: *Adaptive Behavior* 11.4, pp. 209–243.

— (2008). "The Dynamics of Brain-Body-Environment Systems: A Status Report". In: *Handbook of Cognitive Science*. Elsevier Science Ltd, pp. 99–120.

Beer, R. D. and Williams, P. L. (2015). "Information processing and dynamics in minimally cognitive agents". In: *Cognitive science* 39.1, pp. 1–38.

Bellman, R. E. (1957a). "A Markovian decision process". In: *Journal of Mathematics and Mechanics*, pp. 679–684.

— (1957b). *Dynamic Programming*. Courier Dover Publications.

Bernstein, N. (1967). *The co-ordination and regulation of movements*. Oxford: Pergamon Press.

Bialek, W., Nemenman, I., and Tishby, N. (2001). "Predictability, complexity, and learning". In: *Neural computation* 13.11, pp. 2409–2463.

Biehl, M. (2017). "Formal approaches to a definition of agents". In: *arXiv preprint arXiv:1704.02716*.
Biehl, M. et al. (2018). "Expanding the Active Inference Landscape: More Intrinsic Motivations in the Perception-Action Loop". In: *Frontiers in Neurorobotics* 12, p. 45.
Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer-Verlag New York.
Boden, M. A. (2006). *Mind as machine: A history of cognitive science*. Clarendon Press.
Bogacz, R. (2017). "A tutorial on the free-energy framework for modelling perception and learning". In: *Journal of mathematical psychology* 76, pp. 198–211.
Boltzmann, L. (1974). "The second law of thermodynamics". In: *Theoretical physics and philosophical problems*. Springer, pp. 13–32.
Botvinick, M. and Toussaint, M. (2012). "Planning as inference". In: *Trends in cognitive sciences* 16.10, pp. 485–488.
Bowers, J. S. and Davis, C. J. (2012). "Bayesian just-so stories in psychology and neuroscience." In: *Psychological bulletin* 138.3, p. 389.
Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. MIT press.
Briat, C., Gupta, A., and Khammash, M. (2016). "Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks". In: *Cell systems* 2.1, pp. 15–26.
Brooks, R. A. (1986). "A robust layered control system for a mobile robot". In: *Robotics and Automation, IEEE Journal of* 2.1, pp. 14–23.
— (1991a). "Intelligence without representation". In: *Artificial intelligence* 47.1, pp. 139–159.
— (1991b). "New approaches to robotics". In: *Science* 253.5025, pp. 1227–1232.
— (1995). "Intelligence without reason". In: *The artificial life route to artificial intelligence: Building embodied, situated agents*, pp. 25–81.
— (1997). "From earwigs to humans". In: *Robotics and autonomous systems* 20.2, pp. 291–304.
Brown, H. et al. (2013). "Active inference, sensory attenuation and illusions". In: *Cognitive processing* 14.4, pp. 411–427.
Brown, L. D. (1981). "A complete class theorem for statistical problems with finite sample spaces". In: *The Annals of Statistics*, pp. 1289–1300.
Bruineberg, J., Kiverstein, J., and Rietveld, E. (2018). "The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective". In: *Synthese* 195.6, pp. 2417–2444.
Bruineberg, J. et al. (2018). "Free-energy minimization in joint agent-environment systems: A niche construction perspective". In: *Journal of theoretical biology* 455, pp. 161–178.
Buckley, C. L. and Toyoizumi, T. (2018). "A theory of how active behavior stabilises neural activity: Neural gain modulation by closed-loop environmental feedback". In: *PLoS computational biology* 14.1, e1005926.

Buckley, C. L. et al. (2017). "The free energy principle for action and perception: A mathematical review". In: *Journal of Mathematical Psychology* 14, pp. 55–79.

Buhrmann, T. and Di Paolo, E. A. (2014). "Spinal circuits can accommodate interaction torques during multijoint limb movements". In: *Frontiers in computational neuroscience* 8, p. 144.

Buhrmann, T., Di Paolo, E. A., and Barandiaran, X. (2013). "A dynamical systems account of sensorimotor contingencies". In: *Frontiers in psychology* 4, p. 285.

Busse, L. et al. (2017). "Sensation during Active Behaviors". In: *Journal of Neuroscience* 37.45, pp. 10826–10834.

Calvo, P. and Friston, K. J. (2017). "Predicting green: really radical (plant) predictive processing". In: *Journal of The Royal Society Interface* 14.131, p. 20170096.

Carver, C. S. and Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. Springer Science & Business Media.

Chaudhari, P. and Soatto, S. (2018). "Stochastic gradient descent performs variational inference, converges to limit cycles for deep networks". In: *2018 Information Theory and Applications Workshop (ITA)*. IEEE, pp. 1–10.

Chemero, A. (2011). *Radical embodied cognitive science*. MIT press.

Chen, Z. (2003). "Bayesian filtering: From Kalman filters to particle filters, and beyond". In: *Statistics* 182.1, pp. 1–69.

Chevalier, M. et al. (2018). "Design and analysis of a Proportional-Integral-Derivative controller with biological molecules". In: *bioRxiv*. DOI: 10.1101/303545.

Chui, C. K. and Chen, G. (2017). *Kalman filtering with Real-Time Applications*. Springer.

Cisek, P. (1999). "Beyond the computer metaphor: Behaviour as interaction". In: *Journal of Consciousness Studies* 6.11-12, pp. 125–142.

Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT press.

— (2013). "Whatever next? Predictive brains, situated agents, and the future of cognitive science". In: *Behavioral and Brain Sciences* 36.03, pp. 181–204.

— (2015a). "Radical predictive processing". In: *The Southern Journal of Philosophy* 53.S1, pp. 3–27.

— (2015b). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

Clark, A. and Thornton, C. (1997). "Trading spaces: Computation, representation, and the limits of uninformed learning". In: *Behavioral and Brain Sciences* 20.1, pp. 57–66.

Cliff, D., Husbands, P., and Harvey, I. (1993). "Explorations in evolutionary robotics". In: *Adaptive behavior* 2.1, pp. 73–110.

Colombo, M. and Wright, C. (2018). "First principles in the life sciences: the free-energy principle, organicism, and mechanism". In: *Synthese*.

Coltheart, M. (1999). "Modularity and cognition". In: *Trends in cognitive sciences* 3.3, pp. 115–120.

Conant, R. C. and Ashby, W. R. (1970). "Every good regulator of a system must be a model of that system†". In: *International journal of systems science* 1.2, pp. 89–97.

Constant, A. et al. (2018). "A variational approach to niche construction". In: *Journal of The Royal Society Interface* 15.141, p. 20170685.

Cowley, B. et al. (2018). "Precision without Precisions: Handling uncertainty with a single predictive model". In: *Artificial Life Conference Proceedings*. MIT Press, pp. 129–136.

Crapse, T. B. and Sommer, M. A. (2008). "Corollary discharge across the animal kingdom". In: *Nature Reviews Neuroscience* 9.8, p. 587.

Cullen, K. E. (2004). "Sensory signals during active versus passive movement". In: *Current opinion in neurobiology* 14.6, pp. 698–706.

Daunizeau, J. (2017). "The variational Laplace approach to approximate Bayesian inference". In: *arXiv preprint arXiv:1703.02089*.

Dayan, P. et al. (1995). "The Helmholtz Machine". In: *Neural computation* 7.5, pp. 889–904.

De Ridder, D., Vanneste, S., and Freeman, W. (2014). "The Bayesian brain: phantom percepts resolve sensory uncertainty". In: *Neuroscience & Biobehavioral Reviews* 44, pp. 4–15.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38.

Dennett, D. C. (1993). *Consciousness explained*. Penguin uk.

Der, R. and Martius, G. (2012). *The playful machine: theoretical foundation and practical realization of self-organizing robots*. Vol. 15. Springer Science & Business Media.

Der, R., Steinmetz, U., and Pasemann, F. H. (1999). "A new principle to back up evolution with learning". In: *Computational Intelligence for Modelling, Control, and Automation; IOS Press: Amsterdam, Demark*, pp. 43–47.

Di Paolo, E. A. (2005). "Autopoiesis, adaptivity, teleology, agency". In: *Phenomenology and the cognitive sciences* 4.4, pp. 429–452.

Di Paolo, E. A., Buhrmann, T., and Barandiaran, X. (2017). *Sensorimotor Life: An Enactive Proposal*. Oxford University Press.

Di Paolo, E. A. and Iizuka, H. (2008). "How (not) to model autonomous behaviour". In: *BioSystems* 91.2, pp. 409–423.

Doucet, A. and Johansen, A. M. (2009). "A tutorial on particle filtering and smoothing: Fifteen years later". In: *Handbook of nonlinear filtering* 12.656-704, p. 3.

Douven, I. (2017). "Abduction". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University.

Doya, K. (2007). *Bayesian brain: Probabilistic approaches to neural coding*. MIT press.

— (2009). "How can we learn efficiently to act optimally and flexibly?" In: *Proceedings of the National Academy of Sciences* 106.28, pp. 11429–11430.

Engel, A. K., Friston, K. J., and Kragic, D. (2016). "The pragmatic turn: Toward action-oriented views in cognitive science". In:

Ernst, M. O. and Banks, M. S. (2002). "Humans integrate visual and haptic information in a statistically optimal fashion". In: *Nature* 415.6870, p. 429.

Evans, D. J. and Searles, D. J. (2002). "The fluctuation theorem". In: *Advances in Physics* 51.7, pp. 1529–1585.

Faisal, A. A., Selen, L. P., and Wolpert, D. M. (2008). "Noise in the nervous system". In: *Nature reviews neuroscience* 9.4, p. 292.

Feldman, A. G. (2009). "New insights into action–perception coupling". In: *Experimental Brain Research* 194.1, pp. 39–58.

— (2016). "Active sensing without efference copy: referent control of perception". In: *Journal of neurophysiology* 116.3, pp. 960–976.

Feldman, A. G. et al. (2015). "Referent control of action and perception". In: *Challenging Conventional Theories in Behavioral Neuroscience*.

Feldman, H. and Friston, K. J. (2010). "Attention, uncertainty, and free-energy". In: *Frontiers in human neuroscience* 4, p. 215.

Felleman, D. J. and Van, D. E. (1991). "Distributed hierarchical processing in the primate cerebral cortex." In: *Cerebral cortex (New York, NY: 1991)* 1.1, pp. 1–47.

Ferguson, T. S. (1967). *Mathematical statistics: A decision theoretic approach*. Academic press.

Flanders, M. (2011). "What is the biological basis of sensorimotor integration?" In: *Biological cybernetics* 104.1-2, pp. 1–8.

Fodor, J. A. (1983). *The Modularity of Mind*. MIT Press.

— (2001). *The mind doesn't work that way: The scope and limits of computational psychology*. MIT press.

Fox, R. F. (1987). "Stochastic calculus in physics". In: *Journal of statistical physics* 46.5-6, pp. 1145–1157.

Fox, R. and Tishby, N. (2016). "Minimum-information LQG control part I: Memoryless controllers". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, pp. 5610–5616.

Francis, B. A. and Wonham, W. M. (1976). "The internal model principle of control theory". In: *Automatica* 12.5, pp. 457–465.

Franklin, D. W. and Wolpert, D. M. (2011). "Computational mechanisms of sensorimotor control". In: *Neuron* 72.3, pp. 425–442.

Friston, K. J. (2005a). "A theory of cortical responses." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360.1456, pp. 815–836.

— (2005b). "Hallucinations and perceptual inference". In: *Behavioral and Brain Sciences* 28.6, pp. 764–766.

— (2008a). "Hierarchical models in the brain". In: *PLoS Computational Biology* 4.11.

— (2008b). "Variational filtering". In: *NeuroImage* 41.3, pp. 747–766.

— (2009). "The free-energy principle: a rough guide to the brain?" In: *Trends in Cognitive Sciences* 13.7, pp. 293–301.

— (2010a). "Some free-energy puzzles resolved: response to Thornton". In: *Trends in cognitive sciences* 14.2, pp. 54–55.

— (2010b). "The free-energy principle: a unified brain theory?" In: *Nature reviews. Neuroscience* 11.2, pp. 127–138.

Friston, K. J. (2011). "What is optimal about motor control?" In: *Neuron* 72.3, pp. 488–498.

— (2012). "A free energy principle for biological systems". In: *Entropy* 14.11, pp. 2100–2121.

— (2013). "Life as we know it". In: *Journal of the Royal Society Interface* 10.86, p. 20130475.

Friston, K. J., Adams, R. A., and Montague, R. (2012). "What is value? Accumulated reward or evidence?" In: *Frontiers in Neurorobotics* 6.

Friston, K. J. and Ao, P. (2012). "Free energy, value, and attractors". In: *Computational and mathematical methods in medicine* 2012.

Friston, K. J., Daunizeau, J., and Kiebel, S. J. (2009). "Reinforcement learning or active inference?" In: *PLoS One* 4.7, e6421.

Friston, K. J. and Kiebel, S. (2009a). "Attractors in song". In: *New Mathematics and Natural Computation* 5.01, pp. 83–114.

— (2009b). "Cortical circuits for perceptual inference". In: *Neural Networks* 22.8, pp. 1093–1104.

— (2009c). "Predictive coding under the free-energy principle". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 364.1521, pp. 1211–1221.

Friston, K. J., Kilner, J., and Harrison, L. (2006). "A free energy principle for the brain". In: *Journal of Physiology-Paris* 100.1, pp. 70–87.

Friston, K. J., Parr, T., and Vries, B. de (2017). "The graphical brain: belief propagation and active inference". In: *Network Neuroscience* 1.4, pp. 381–414.

Friston, K. J., Samothrakis, S., and Montague, R. (2012a). "Active inference and agency: Optimal control without cost functions". In: *Biological Cybernetics* 106.8-9, pp. 523–541.

— (2012b). "Active inference and agency: Optimal control without cost functions". In: *Biological Cybernetics* 106.8-9, pp. 523–541.

Friston, K. J., Thornton, C., and Clark, A. (2012). "Free-energy minimization and the dark-room problem". In: *Frontiers in psychology* 3, p. 130.

Friston, K. J., Trujillo-Barreto, N., and Daunizeau, J. (2008). "DEM: A variational treatment of dynamic systems". In: *NeuroImage* 41.3, pp. 849–885.

Friston, K. J. et al. (2007). "Variational free energy and the Laplace approximation". In: *Neuroimage* 34.1, pp. 220–234.

Friston, K. J. et al. (2010a). "Action and behavior: A free-energy formulation". In: *Biological Cybernetics* 102.3, pp. 227–260.

Friston, K. J. et al. (2010b). "Generalised filtering". In: *Mathematical Problems in Engineering* 2010.

Friston, K. J. et al. (2012). "Perceptions as hypotheses: saccades as experiments". In: *Frontiers in psychology* 3, p. 151.

Friston, K. J. et al. (2013). "The anatomy of choice: active inference and agency." In: *Frontiers in human neuroscience* 7.September, p. 598.

Friston, K. J. et al. (2015). "Active inference and epistemic value." In: *Cognitive neuroscience*, pp. 1–28.

Friston, K. J. et al. (2016a). "Active inference and learning". In: *Neuroscience & Biobehavioral Reviews* 68, pp. 862–879.

Friston, K. J. et al. (2016b). "The dysconnection hypothesis (2016)". In: *Schizophrenia Research* 176.2, pp. 83–94.

Friston, K. J. et al. (2017). "Active inference: a process theory". In: *Neural Computation* 29.1, pp. 1–49.

Froese, T. (2010). "From Cybernetics to Second-Order Cybernetics: A Comparative Analysis of Their Central Ideas." In: *Constructivist Foundations* 5.2.

— (2011). "From Second-order Cybernetics to Enactive Cognitive Science: Varela's Turn From Epistemology to Phenomenology". In: *Systems Research and Behavioral Science* 28.6, pp. 631–645.

Froese, T. and Ikegami, T. (2013). "The brain is not an isolated "black box," nor is its goal to become one". In: *Behavioral and Brain Sciences* 36.3, pp. 213–214.

Froese, T. and Stewart, J. (2010). "Life after Ashby: ultrastability and the autopoietic foundations of biological autonomy". In: *Cybernetics & Human Knowing* 17.4, pp. 7–49.

Froese, T. and Ziemke, T. (2009). "Enactive artificial intelligence: Investigating the systemic organization of life and mind". In: *Artificial Intelligence* 173.3-4, pp. 466–500.

Gallagher, S. (2006). *How the body shapes the mind*. Clarendon Press.

Garpinger, O., Hägglund, T., and Åström, K. J. (2014). "Performance and robustness trade-offs in PID control". In: *Journal of Process Control* 24.5, pp. 568–577.

Gastpar, M., Rimoldi, B., and Vetterli, M. (2003). "To code, or not to code: Lossy source-channel communication revisited". In: *IEEE Transactions on Information Theory* 49.5, pp. 1147–1158.

George, D. and Hawkins, J. (2009). "Towards a mathematical theory of cortical microcircuits". In: *PLoS computational biology* 5.10, e1000532.

Georgiou, T. T. and Lindquist, A. (2013). "The separation principle in stochastic control, redux". In: *IEEE Transactions on Automatic Control* 58.10, pp. 2481–2494.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton, Mifflin and Company.

Gładziejewski, P. (2016). "Predictive coding and representationalism". In: *Synthese* 193.2, pp. 559–582.

Gregory, R. L. (1980). "Perceptions as hypotheses". In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 290.1038, pp. 181–197.

Gregory, R. L. (1970). *The intelligent eye.* ERIC.

Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (2008). "Bayesian models of cognition". In:

Griffiths, T. L. et al. (2010). "Probabilistic models of cognition: Exploring representations and inductive biases". In: *Trends in cognitive sciences* 14.8, pp. 357–364.

Hägglund, T. (1995). "A control-loop performance monitor". In: *Control Engineering Practice* 3.11, pp. 1543–1551.

Harvey, I. et al. (2005). "Evolutionary robotics: A new scientific tool for studying cognition". In: *Artificial life* 11.1-2, pp. 79–98.

Helmholtz, H. von (1867). *Handbuch der physiologischen Optik*. Vol. 9. Voss.

Hinton, G. E. and Zemel, R. S. (1994). "Autoencoders, minimum description length and Helmholtz free energy". In: *Advances in neural information processing systems*, pp. 3–10.

Hinton, G. E. et al. (1995). "The" wake-sleep" algorithm for unsupervised neural networks". In: *Science* 268.5214, pp. 1158–1161.

Ho, B. and Kalman, R. E. (1966). "Effective construction of linear state-variable models from input/output functions". In: *at-Automatisierungstechnik* 14.1-12, pp. 545–548.

Hohwy, J. (2013). *The predictive mind*. OUP Oxford.

— (2014). "The Self-Evidencing Brain". In: *Noûs*.

Holst, E. von and Mittelstaedt, H. (1950). "Das reafferenzprinzip". In: *Naturwissenschaften* 37.20, pp. 464–476.

Hosoya, T., Baccus, S. A., and Meister, M. (2005). "Dynamic predictive coding by the retina". In: *Nature* 436.7047, p. 71.

Huang, Y. and Rao, R. P. (2011). "Predictive coding". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 2.5, pp. 580–593.

Hurley, S. (2001). "Perception and action: Alternative views". In: *Synthese* 129.1, pp. 3–40.

Huys, Q. J., Maia, T. V., and Frank, M. J. (2016). "Computational psychiatry as a bridge from neuroscience to clinical applications". In: *Nature neuroscience* 19.3, p. 404.

Iizuka, H. and Ikegami, T. (2004). "Simulating autonomous coupling in discrimination of light frequencies". In: *Connection Science* 16.4, pp. 283–299.

Isomura, T. and Friston, K. (2018). "In vitro neural networks minimise variational free energy". In: *Scientific reports* 8.1.

Isomura, T., Kotani, K., and Jimbo, Y. (2015). "Cultured cortical neurons can perform blind source separation according to the free-energy principle". In: *PLoS computational biology* 11.12, e1004643.

Itti, L. and Baldi, P. (2009). "Bayesian surprise attracts human attention". In: *Vision research* 49.10, pp. 1295–1306.

Jaynes, E. T. (1957a). "Information theory and statistical mechanics". In: *Physical review* 106.4, p. 620.

— (1957b). "Information theory and statistical mechanics. II". In: *Physical review* 108.2, p. 171.

— (2003). *Probability theory: The logic of science*. Cambridge university press.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Vol. 64. Academic Press.

Johnson, M. A. and Moradi, M. H. (2005). *PID control*. Springer.

Jones, M. and Love, B. C. (2011). “Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition”. In: *Behavioral and Brain Sciences* 34.4, pp. 169–188.

Jordan, M. I. (1996). “Computational aspects of motor control and motor learning”. In: *Handbook of perception and action*. Vol. 2. Elsevier, pp. 71–120.

Jordan, R., Kinderlehrer, D., and Otto, F. (1998). “The variational formulation of the Fokker–Planck equation”. In: *SIAM journal on mathematical analysis* 29.1, pp. 1–17.

Jung, T., Polani, D., and Stone, P. (2011). “Empowerment for continuous agent—environment systems”. In: *Adaptive Behavior* 19.1, pp. 16–39.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). “Reinforcement learning: A survey”. In: *Journal of artificial intelligence research* 4, pp. 237–285.

Kalman, R. E. (1960a). “On the general theory of control systems”. In: *Proceedings First International Conference on Automatic Control, Moscow, USSR*.

Kalman, R. E. (1960b). “A new approach to linear filtering and prediction problems”. In: *Journal of basic Engineering* 82.1, pp. 35–45.

— (1960c). “Contributions to the theory of optimal control”. In: *Bol. Soc. Mat. Mexicana* 5.2, pp. 102–119.

Kalman, R. E. and Bucy, R. S. (1961). “New results in linear filtering and prediction theory”. In: *Journal of basic engineering* 83.1, pp. 95–108.

Kappen, H. J. (2005). “Linear theory for control of nonlinear stochastic systems”. In: *Physical review letters* 95.20, p. 200201.

— (2011). “Optimal control theory and the linear Bellman equation”. In:

Kappen, H. J., Gómez, V., and Opper, M. (2012). “Optimal control as a graphical model inference problem”. In: *Machine learning* 87.2, pp. 159–182.

Kawato, M. (1999). “Internal models for motor control and trajectory planning”. In: *Current opinion in neurobiology* 9.6, pp. 718–727.

Keller, G. B. and Mrsic-Flogel, T. D. (2018). “Predictive Processing: A Canonical Cortical Computation”. In: *Neuron* 100.2, pp. 424–435.

Kiebel, S. J., Daunizeau, J., and Friston, K. J. (2008). “A hierarchy of time-scales and the brain”. In: *PLoS Comput Biol* 4.11, e1000209.

Kim, C. S. (2018). “Recognition Dynamics in the Brain under the Free-Energy Principle”. In: *Neural Computation* 30.10, pp. 2616–2659.

Kirchhoff, M. et al. (2018). “The Markov blankets of life: autonomy, active inference and the free energy principle”. In: *Journal of The Royal Society Interface* 15.138, p. 20170792.

Kirchhoff, M. D. (2018). “Autopoiesis, free energy, and the life–mind continuity thesis”. In: *Synthese* 195.6, pp. 2519–2540.

Kirchhoff, M. D. and Froese, T. (2017). “Where there is life there is mind: In support of a strong life-mind continuity thesis”. In: *Entropy* 19.4, p. 169.

Kirsh, D. (1991). “Today the earwig, tomorrow man?” In: *Artificial intelligence* 47.1-3, pp. 161–184.

Klöden, P. E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. Springer.

Klyubin, A. S., Polani, D., and Nehaniv, C. L. (2005a). "All else being equal be empowered". In: *European Conference on Artificial Life*. Springer, pp. 744–753.

— (2005b). "Empowerment: A universal agent-centric measure of control". In: *Evolutionary Computation, 2005. The 2005 IEEE Congress on*. Vol. 1. IEEE, pp. 128–135.

Kneller, A. and Thornton, J. (2015). "Distal dendrite feedback in hierarchical temporal memory". In: *Neural Networks (IJCNN), 2015 International Joint Conference on*. IEEE, pp. 1–8.

Knill, D. C. and Pouget, A. (2004). "The Bayesian brain: the role of uncertainty in neural coding and computation". In: *Trends in Neurosciences* 27.12, pp. 712–719.

Knill, D. C. and Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.

Kolchinsky, A. and Wolpert, D. H. (2018). "Semantic information, autonomous agency and non-equilibrium statistical physics". In: *Interface Focus* 8.6, p. 20180041.

Körding, K. P. (2007). "Decision theory: what" should" the nervous system do?" In: *Science* 318.5850, pp. 606–610.

Körding, K. P. and Wolpert, D. M. (2006). "Bayesian decision theory in sensorimotor control". In: *Trends in cognitive sciences* 10.7, pp. 319–326.

Körding, K. P. and Wolpert, D. M. (2004). "The loss function of sensorimotor learning". In: *Proceedings of the National Academy of Sciences* 101.26, pp. 9839–9842.

Krakauer, J. W. et al. (2017). "Neuroscience needs behavior: correcting a reductionist Bias". In: *Neuron* 93.3, pp. 480–490.

Kullback, S. and Leibler, R. A. (1951). "On information and sufficiency". In: *The annals of mathematical statistics* 22.1, pp. 79–86.

Lee, T. S. and Mumford, D. (2003). "Hierarchical Bayesian inference in the visual cortex". In: *JOSA A* 20.7, pp. 1434–1448.

Li, B. et al. (2011). "Generalised filtering and stochastic DCM for fMRI". In: *neuroimage* 58.2, pp. 442–457.

Li, W. and Todorov, E. (2004). "Iterative linear quadratic regulator design for nonlinear biological movement systems." In: *ICINCO (1)*, pp. 222–229.

Linson, A. et al. (2018). "The active inference approach to ecological perception: General information dynamics for natural and artificial embodied Cognition". In: *Frontiers in Robotics and AI* 5, p. 21.

Little, D. Y.-J. and Sommer, F. T. (2013). "Maximal mutual information, not minimal entropy, for escaping the "Dark Room"". In: *Behavioral and Brain Sciences* 36.3, pp. 220–221.

Loeb, G. E. (2012). "Optimal isn't good enough". In: *Biological cybernetics* 106.11-12, pp. 757–765.

Longtin, A. (2003). "Effects of noise on nonlinear dynamics". In: *Nonlinear Dynamics in Physiology and Medicine*. Springer, pp. 149–189.

Longtin, A. (2010). “Stochastic dynamical systems”. In: *Scholarpedia* 5.4. revision #124121, p. 1619.

Łuczka, J (2005). “Non-Markovian stochastic processes: Colored noise”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 15.2, p. 026107.

MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

Makhoul, J. (1975). “Linear prediction: A tutorial review”. In: *Proceedings of the IEEE* 63.4, pp. 561–580.

Mandt, S., Hoffman, M., and Blei, D. (2016). “A variational analysis of stochastic gradient algorithms”. In: *International Conference on Machine Learning*, pp. 354–363.

Marr, D. (1982). *Vision: A computational approach*. Freeman.

Maturana, H. R. (2011). “Ultrastability… autopoiesis? Reflective response to Tom Froese and John Stewart”. In: *Cybernetics & Human Knowing* 18.1-2, pp. 143–152.

Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Springer Science & Business Media.

Maxwell, J. C. (1868). “I. On governors”. In: *Proceedings of the Royal Society of London* 16, pp. 270–283.

McCall, R. and Franklin, S. (2013). “Cortical learning algorithms with predictive coding for a systems-level cognitive architecture”. In: *Second annual conference on advances in cognitive systems poster collection*, pp. 149–66.

McCulloch, W. S. and Pitts, W. (1943). “A logical calculus of the ideas immanent in nervous activity”. In: *The bulletin of mathematical biophysics* 5.4, pp. 115–133.

McGregor, S. (2017). “The Bayesian stance: Equations for ‘as-if’sensorimotor agency”. In: *Adaptive Behavior* 25.2, pp. 72–82.

McGregor, S., Baltieri, M., and Buckley, C. L. (2015). “A Minimal Active Inference Agent”. In: *arXiv preprint arXiv:1503.04187*.

Meinhold, R. J. and Singpurwalla, N. D. (1983). “Understanding the Kalman filter”. In: *The American Statistician* 37.2, pp. 123–127.

Minka, T. P. (2001). “Expectation propagation for approximate Bayesian inference”. In: *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., pp. 362–369.

Mitter, S. K. and Newton, N. J. (2003). “A variational approach to nonlinear estimation”. In: *SIAM journal on control and optimization* 42.5, pp. 1813–1833.

Montague, P. R. et al. (2012). “Computational psychiatry”. In: *Trends in cognitive sciences* 16.1, pp. 72–80.

Mumford, D. (1992). “On the computational architecture of the neocortex”. In: *Biological cybernetics* 66.3, pp. 241–251.

Muzzey, D. et al. (2009). “A systems-level analysis of perfect adaptation in yeast osmoregulation”. In: *Cell* 138.1, pp. 160–171.

Nagengast, A. J., Braun, D. A., and Wolpert, D. M. (2010). "Risk-sensitive optimal feedback control accounts for sensorimotor behavior under uncertainty". In: *PLoS computational biology* 6.7, e1000857.

Najafi, F. and Churchland, A. K. (2018). "Perceptual Decision-Making: A Field in the Midst of a Transformation". In: *Neuron* 100.2, pp. 453–462.

Neisser, U. (1967). *Cognitive psychology.* Appleton-Century-Crofts.

Neumann, J. von (2012). *The computer and the brain*. Yale University Press.

Newell, A., Simon, H. A., et al. (1972). *Human problem solving*. Vol. 104. 9. Prentice-Hall Englewood Cliffs, NJ.

Newen, A., De Bruin, L., and Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. Oxford University Press.

Nolfi, S., Floreano, D., and Floreano, D. D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. MIT press.

Odling-Smee, F. J., Laland, K. N., and Feldman, M. W. (2003). *Niche construction: the neglected process in evolution*. 37. Princeton university press.

Opper, M. and Archambeau, C. (2009). "The variational Gaussian approximation revisited". In: *Neural computation* 21.3, pp. 786–792.

Ozaki, T. (1992). "A bridge between nonlinear time series models and nonlinear stochastic dynamical systems: a local linearization approach". In: *Statistica Sinica*, pp. 113–135.

Parr, T. and Friston, K. J. (2017). "Uncertainty, epistemics and active inference". In: *Journal of The Royal Society Interface* 14.136, p. 20170376.

— (2018a). "Generalised free energy and active inference: can the future cause the past?" In: *bioRxiv*.

— (2018b). "The anatomy of inference: Generative models and brain structure". In: *Frontiers in computational neuroscience* 12.

— (2018c). "The Discrete and Continuous Brain: From Decisions to Movement—and Back Again". In: *Neural computation* Early Access, pp. 1–29.

Parrondo, J. M., Horowitz, J. M., and Sagawa, T. (2015). "Thermodynamics of information". In: *Nature physics* 11.2, p. 131.

Pearl, J. (2001). "Bayesianism and causality, or, why I am only a half-Bayesian". In: *Foundations of bayesianism*. Springer, pp. 19–36.

— (2009). "Causal inference in statistics: An overview". In: *Statistics Surveys* 3, pp. 96–146.

Pearl, J. and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. 1st. New York, NY, USA: Basic Books, Inc.

Penny, W., Kiebel, S., and Friston, K. J. (2006). "Variational Bayes". In: *Statistical Parametric Mapping: The analysis of functional brain images*. Ed. by K. S. N. T. P. W. Friston K Ashburner J. Elsevier. Chap. 24, pp. 303–312.

Penny, W. (2012). "Bayesian models of brain and behaviour". In: *ISRN Biomathematics* 2012.

Perrinet, L. U., Adams, R. A., and Friston, K. J. (2014). "Active inference, eye movements and oculomotor delays". In: *Biological cybernetics* 108.6, pp. 777–801.

Pezzulo, G., Rigoli, F., and Friston, K. J. (2015). "Active inference, homeostatic regulation and adaptive behavioural control". In: *Progress in neurobiology* 134, pp. 17–35.

Pezzulo, G. et al. (2017). "Model-Based Approaches to Active Perception and Control". In: *Entropy* 19.6, p. 266.

Pfeifer, R. and Scheier, C. (2001). *Understanding intelligence*. MIT press.

Pickering, M. J. and Clark, A. (2014). "Getting ahead: forward models and their place in cognitive architecture". In: *Trends in cognitive sciences* 18.9, pp. 451–456.

Pontryagin, L. S. et al. (1962). *The mathematical theory of optimal processes*. Wiley.

Powers, W. T. (1973). *Behavior: The control of perception*. Aldine Chicago.

Prinz, J. J. (2006). "Is the Mind Really Modular?" In: *Contemporary Debates in Cognitive Science*. Ed. by R. J. Stainton. Blackwell, pp. 22–36.

Rahnev, D. and Denison, R. N. (2018). "Suboptimality in Perceptual Decision Making". In: *Behavioral and Brain Sciences*, pp. 1–107.

Ramstead, M. J. D., Badcock, P. B., and Friston, K. J. (2017). "Answering Schrödinger's question: a free-energy formulation". In: *Physics of life reviews*.

Rao, R. P. (1999). "An optimal estimation approach to visual perception and learning". In: *Vision research* 39.11, pp. 1963–1989.

Rao, R. P. and Ballard, D. H. (1999). "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects". In: *Nature neuroscience* 2.1, pp. 79–87.

Rao, R. P. et al. (2002). *Probabilistic models of the brain: Perception and neural function*. MIT press.

Rao, V. G. and Bernstein, D. S. (2001). "Naive control of the double integrator". In: *IEEE Control Systems* 21.5, pp. 86–97.

Redish, A. D. (2004). "Addiction as a computational process gone awry". In: *Science* 306.5703, pp. 1944–1947.

Redish, A. D. and Gordon, J. A. (2016). *Computational psychiatry: New perspectives on mental illness*. MIT Press.

Rescorla, R. A., Wagner, A. R., et al. (1972). "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement". In: *Classical conditioning II: Current research and theory* 2, pp. 64–99.

Ritz, H. et al. (2018). "A Control Theoretic Model of Adaptive Learning in Dynamic Environments". In: *Journal of cognitive neuroscience*, pp. 1–17.

Rivera, D. E., Morari, M., and Skogestad, S. (1986). "Internal model control: PID controller design". In: *Industrial & engineering chemistry process design and development* 25.1, pp. 252–265.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media.

Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). "Behavior, purpose and teleology". In: *Philosophy of science* 10.1, pp. 18–24.

Roweis, S. and Ghahramani, Z. (2001). "Learning nonlinear dynamical systems using the Expectation-Maximization algorithm". In: *Kalman filtering and neural networks*. Ed. by S. Haykin. Wiley Online Library. Chap. 10, pp. 175–220.

Russell, S. and Norvig, P. (2009). *Artificial Intelligence: A Modern Approach*. 3rd. Upper Saddle River, NJ, USA: Prentice Hall Press.

Salge, C., Glackin, C., and Polani, D. (2014). "Empowerment – An introduction". In: *Guided Self-Organization: Inception*. Springer, pp. 67–114.

Sanborn, A. N. and Chater, N. (2016). "Bayesian brains without probabilities". In: *Trends in cognitive sciences* 20.12, pp. 883–893.

Särkkä, S. (2013). *Bayesian filtering and smoothing*. Vol. 3. Cambridge University Press.

Schaal, S., Mohajerian, P., and Ijspeert, A. (2007). "Dynamics systems vs. optimal control — a unifying view". In: *Progress in brain research* 165, pp. 425–445.

Schrödinger, E. (1944). *What Is Life? the physical aspect of the living cell and mind*. Cambridge University Press, Cambridge.

Schwartenbeck, P. et al. (2015). "Optimal inference with suboptimal models: addiction and active Bayesian inference". In: *Medical hypotheses* 84.2, pp. 109–117.

Schwartz, A. B. (2016). "Movement: how the brain communicates with the world". In: *Cell* 164.6, pp. 1122–1135.

Scott, S. H. (2004). "Optimal feedback control and the neural basis of volitional motor control". In: *Nature Reviews Neuroscience* 5.7, p. 532.

Seifert, U. (2012). "Stochastic thermodynamics, fluctuation theorems and molecular machines". In: *Reports on progress in physics* 75.12, p. 126001.

Sengupta, B., Stemmler, M. B., and Friston, K. J. (2013). "Information and efficiency in the nervous system? A synthesis". In: *PLoS Comput Biol* 9.7, e1003157.

Seth, A. K. (2014a). "A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia". In: *Cognitive neuroscience* 5.2, pp. 97–118.

— (2014b). "The Cybernetic Bayesian Brain". In: *Open MIND*. Open MIND. Frankfurt am Main: MIND Group.

Seth, A. K. and Tsakiris, M. (2018). "Being a beast machine: The somatic basis of selfhood". In: *Trends in cognitive sciences*.

Shannon, C. E. (1948). "A mathematical theory of communication". In: *Bell system technical journal* 27.3, pp. 379–423.

— (1959). "Coding theorems for a discrete source with a fidelity criterion". In: *IRE Nat. Conv. Rec* 4.142-163, p. 1.

Shipp, S., Adams, R. A., and Friston, K. J. (2013). "Reflections on agranular architecture: predictive coding in the motor cortex". In: *Trends in neurosciences* 36.12, pp. 706–716.

Simon, H. A. (1956). "Dynamic programming under uncertainty with a quadratic criterion function". In: *Econometrica, Journal of the Econometric Society*, pp. 74–81.

Simon, H. A. (1991). "The architecture of complexity". In: *Facets of systems science*. Springer, pp. 457–476.

Simon, H. A. and Ando, A. (1961). "Aggregation of variables in dynamic systems". In: *Econometrica (pre-1986)* 29.2, p. 111.

Sontag, E. D. (2003). "Adaptation and regulation with signal detection implies internal model". In: *Systems & control letters* 50.2, pp. 119–126.

Sorenson, H. W. (1970). "Least-squares estimation: from Gauss to Kalman". In: *IEEE spectrum* 7.7, pp. 63–68.

Sporns, O. and Edelman, G. M. (1993). "Solving Bernstein's problem: A proposal for the development of coordinated movement by selection". In: *Child development* 64.4, pp. 960–981.

Spratling, M. W. (2008). "Predictive coding as a model of biased competition in visual attention". In: *Vision research* 48.12, pp. 1391–1408.

— (2013). "Distinguishing theory from implementation in predictive coding accounts of brain function". In: *Behavioral and Brain Sciences* 36.3, pp. 231–232.

— (2017). "A review of predictive coding algorithms". In: *Brain and cognition* 112, pp. 92–97.

Spratling, M. (2016). "Predictive coding as a model of cognition". In: *Cognitive processing*, pp. 1–27.

Srinivasan, M. V., Laughlin, S. B., and Dubs, A. (1982). "Predictive coding: a fresh view of inhibition in the retina". In: *Proceedings of the Royal Society of London B: Biological Sciences* 216.1205, pp. 427–459.

Stengel, R. F. (1994). *Optimal control and estimation*. Courier Corporation.

Stephan, K. E. and Mathys, C. (2014). "Computational approaches to psychiatry". In: *Current opinion in neurobiology* 25, pp. 85–92.

Stevenson, I. H. et al. (2009). "Bayesian integration and non-linear feedback control in a full-body motor task". In: *PLoS computational biology* 5.12, e1000629.

Still, S. et al. (2012). "Thermodynamics of prediction". In: *Physical review letters* 109.12, p. 120604.

Straka, H., Simmers, J., and Chagnaud, B. P. (2018). "A New Perspective on Predictive Motor Signaling". In: *Current Biology* 28.5, R232–R243.

Stratonovich, R. L. (1967). *Topics in the theory of random noise*. Vol. 2. CRC Press.

Sussmann, H. J. and Willems, J. C. (1997). "300 years of optimal control: from the brachystochrone to the maximum principle". In: *IEEE Control Systems* 17.3, pp. 32–44.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT press.

Svrcek, W. Y., Mahoney, D. P., Young, B. R., et al. (2006). *A real-time approach to process control*. Wiley New York.

Tanaka, T., Esfahani, P. M., and Mitter, S. K. (2015). "LQG control with minimal information: Three-stage separation principle and SDP-based solution synthesis". In: *arXiv preprint arXiv:1510.04214*.

Tatikonda, S. C. (2000). "Control under communication constraints". PhD thesis. Massachusetts Institute of Technology.

Theil, H. (1957). "A note on certainty equivalence in dynamic planning". In: *Econometrica: Journal of the Econometric Society*, pp. 346–349.

Thornton, C. (2014). "Infotropism as the underlying principle of perceptual organization". In: *Journal of Mathematical Psychology* 61, pp. 38–44.

— (2016). "Predictive processing simplified: The infotropic machine." In: *Brain and cognition*.

Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic robotics*. MIT press.

Tishby, N., Pereira, F. C., and Bialek, W. (1999). "The information bottleneck method". In: *Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing*. University of Illinois Press, pp. 368–377.

Tishby, N. and Polani, D. (2011). "Information theory of decisions and actions". In: *Perception-action cycle*. Springer, pp. 601–636.

Todorov, E. (2004). "Optimality principles in sensorimotor control". In: *Nature neuroscience* 7.9, pp. 907–915.

— (2005). "Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system". In: *Neural computation* 17.5, pp. 1084–1108.

— (2006). "Optimal control theory". In: *Bayesian brain: probabilistic approaches to neural coding*, pp. 269–298.

— (2008). "General duality between optimal control and estimation". In: *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, pp. 4286–4292.

— (2009a). "Efficient computation of optimal actions". In: *Proceedings of the national academy of sciences* 106.28, pp. 11478–11483.

— (2009b). "Parallels between sensory and motor information processing". In: *The cognitive neurosciences,* pp. 613–24.

Todorov, E. and Jordan, M. I. (2002). "Optimal feedback control as a theory of motor coordination". In: *Nature neuroscience* 5.11, p. 1226.

Tribus, M. (1961). *Thermostatics and thermodynamics: an introduction to energy, information and states of matter, with engineering applications*. van Nostrand.

Turing, A. M. (1937). "On computable numbers, with an application to the Entscheidungsproblem". In: *Proceedings of the London mathematical society* 2.1, pp. 230–265.

— (1950). "Computing Machinery and Intelligence". In: *Mind* 49, pp. 433–460.

Valdes-Sosa, P. A. et al. (2011). "Effective connectivity: influence, causality and biophysical modeling". In: *Neuroimage* 58.2, pp. 339–361.

Van Gelder, T. (1995). "What might cognition be, if not computation?" In: *The Journal of Philosophy* 92.7, pp. 345–381.

— (1998). "The dynamical hypothesis in cognitive science". In: *Behavioral and brain sciences* 21.5, pp. 615–628.

Van Kampen, N. G. (1981). "Itô versus stratonovich". In: *Journal of Statistical Physics* 24.1, pp. 175–187.

Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*. Vol. 1. Elsevier.

Van Overschee, P. and De Moor, B. (1993). "Subspace algorithms for the stochastic identification problem". In: *Automatica* 29.3, pp. 649–660.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

Velichkovsky, B. M. (2005). "Modularity of cognitive organization: why it is so appealing and why it is wrong". In: *Modularity: Understanding the development and evolution of natural complex systems*. Ed. by W. Callebaut and D. Rasskin-Gutman. MIT Press Cambridge, MA, pp. 335–356.

Villalobos, M. (2013). "Enactive cognitive science: revisionism or revolution?" In: *Adaptive Behavior* 21.3, pp. 159–167.

Wainwright, M. J., Jordan, M. I., et al. (2008). "Graphical models, exponential families, and variational inference". In: *Foundations and Trends in Machine Learning* 1.1–2, pp. 1–305.

Walter, W. G. (1950). "An imitation of life". In: *Scientific American* 182.5, pp. 42–45.

Weber, A. and Varela, F. J. (2002). "Life after Kant: Natural purposes and the autopoietic foundations of biological individuality". In: *Phenomenology and the cognitive sciences* 1.2, pp. 97–125.

Whittle, P. (1981). "Risk-sensitive linear/quadratic/Gaussian control". In: *Advances in Applied Probability* 13.4, pp. 764–777.

Widrow, B. and Hoff, M. E. (1960). *Adaptive switching circuits*. Tech. rep. Stanford Univ Ca Stanford Electronics Labs.

Wiener, N. (1961). *Cybernetics or Control and Communication in the Animal and the Machine*. Vol. 25. MIT press.

Wiese, W. (2016). "Action Is Enabled by Systematic Misrepresentations". In: *Erkenntnis*, pp. 1–20.

Wilson, M. (2002). "Six views of embodied cognition". In: *Psychonomic bulletin & review* 9.4, pp. 625–636.

Wilson, R. A. and Foglia, L. (2017). "Embodied Cognition". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Spring 2017. Metaphysics Research Lab, Stanford University.

Wittenmark, B. (1995). "Adaptive dual control methods: An overview". In: *Adaptive Systems in Control and Signal Processing 1995*. Elsevier, pp. 67–72.

Wolpert, D. M. (1997). "Computational approaches to motor control". In: *Trends in cognitive sciences* 1.6, pp. 209–216.

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). "Principles of sensorimotor learning". In: *Nature Reviews Neuroscience* 12.12, p. 739.

Wolpert, D. M. and Ghahramani, Z. (2000). "Computational principles of movement neuroscience". In: *Nature neuroscience* 3.11s, p. 1212.

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). "An internal model for sensorimotor integration". In: *Science* 269.5232, pp. 1880–1882.

Wonham, W. (1968). "On the separation theorem of stochastic control". In: *SIAM Journal on Control* 6.2, pp. 312–326.

Yang, L. and Iglesias, P. A. (2006). "Positive feedback may cause the biphasic response observed in the chemoattractant-induced response of Dictyostelium cells". In: *Systems & control letters* 55.4, pp. 329–337.

Yang, S. C.-H., Wolpert, D. M., and Lengyel, M. (2016). "Theoretical perspectives on active sensing". In: *Current Opinion in Behavioral Sciences* 11, pp. 100–108.

Yeo, S.-H., Franklin, D. W., and Wolpert, D. M. (2016). "When optimal feedback control is not enough: Feedforward strategies are required for optimal control with active sensing". In: *PLoS computational biology* 12.12, e1005190.

Yi, T.-M. et al. (2000). "Robust perfect adaptation in bacterial chemotaxis through integral feedback control". In: *Proceedings of the National Academy of Sciences* 97.9, pp. 4649–4653.

Yuille, A. and Kersten, D. (2006). "Vision as Bayesian inference: analysis by synthesis?" In: *Trends in cognitive sciences* 10.7, pp. 301–308.

Zahavi, D. (2017). "Brain, Mind, World: Predictive coding, neo-Kantianism, and transcendental idealism". In: *Husserl Studies*, pp. 1–15.

Zelik, K. E. et al. (2014). "Can modular strategies simplify neural control of multidirectional human locomotion?" In: *Journal of neurophysiology* 111.8, pp. 1686–1702.